الحاسوب واللغة العربية

تطبيقات عملية

(تعليم، ترجمة، صناعة المعجم، بحث)

الرياض

١٤٤٢هـ/ ٢٠٢١

المقدمة

الحمد لله خلق الإنسان علمه البيان، ومن آياته اختلاف الألسنة والألوان. أنزل القرآن بلسان عربي مبين. والصلاة والسلام على أفصح من تكلم بلغة القرآن. وبعد:

لقد مضت عقود وأنا أعايش أو أتابع المجالات التي تطور فيها الحاسوب في خدمة اللغة عامة والعربية خاصة، وأمضيت أوقاتًا مديدة (منها سنوات التفرغ العلمي) وأنا أدرس هذه المجالات المختلفة سواء في تعليم اللغة العربية أم في تطوير أدوات البحث فيها أم الترجمة منها وإليها، إضافة إلى التطورات المتسارعة في تطوير البرمجيات المساعدة في خدمة اللغة العربية. كما شرفت بإنشاء بنك آلي للمصطلحات لخدمة اللغة العربية والناطقين بها (البنك الآلي السعودي للمصطلحات –باسم) في مدينة الملك عبدالعزيز للعلوم والتقنية بالرياض.

وقد تجمعت لدي مجموعة من الدراسات والأبحاث في هذه المجالات المختلفة، نشرت بعضها أو قدمته في مؤتمرات أو ندوات علمية. فرأيت تعميمًا للفائدة للقارئ العربي أن أختار أهمها وأقدمها له، بعد تنقيحها وتحديث المعلومات فيها. لذلك قد يجد القارئ بعضًا من التكرار أحيانًا.

وها هي أقدمها بين دفتي هذا الكتاب الذي أسميته: "الحاسوب واللغة العربية: تطبيقات عملية. وقد راعيت في كل هذه الدراسات أن يكون الحديث موجهًا للمثقفين العرب من معلمين ومترجمين وباحثين، وللساني العادي، غير المتخصص في علوم الحاسوب والمعلوماتية، ولجأت إلى تبسيط الكثير من المصطلحات الفنية، كي تعم الفائدة لأكبر عدد من القراء.

يبدأ الكتاب بمدخل عام عن علاقة الحاسوب باللغة، مع نماذج لهذه العلاقة ومجالات دراستها. ويختتم بعرض لأهم أساليب الاستفادة من الحاسوب في خدمة اللغة العربية.

وقد قُسِّم الكتاب إلى أربعة أقسام رئيسة هي:

القسم الأول: التعليم والتقويم، حيث نجد مناقشة لمميزات الحاسوب في التعليم، ثم كيفية الاستفادة منه في تعليم عناصر اللغة من أصوات ومفردات وتراكيب، متبوعًا بتعليم مهارات الاستماع والقراءة والكلام والكتابة. يلي ذلك مناقشة لأساليب الاستفادة من الحاسوب في التقويم اللغوي. ونختم القسم بأهم معوقات استخدام الحاسوب في تعليم اللغة مع اقتراحات لبعض الحلول المناسبة.

القسم الثاني: الترجمة الآلية، وفيه ثلاثة فصول: نتحدث في الأول منها عن مفهوم الترجمة الآلية والترجمة بمعاونة الحاسوب، إضافة إلى المعينات الحاسوبية في الترجمة، متبوعًا باستعراض لأهم

البرامج الحاسوبية للترجمة الآلية إلى العربية والمتاحة في صور مختلفة: على الحواسيب المستقلة وعلى ا لشابكة (الانترنت) والهواتف الذكية. كما نعرض بإيجاز لبرمجيات الترجمة الآلية للغة المنطوقة. ونختتم الفصل بمناقشة لأهم مشكلات الترجمة الآلية، إضافة إلى أساليب تقويم الترجمة الآلية ومخرجاتها.

أما الفصل الثاني، فقد خصص للحديث عن أهم المشكلات التي تواجه الترجمة الآلية من العربية إلى الإنجليزية أو العكس، مع نماذج تطبيقية، وأسباب ذلك. ونختم بالإشارة بإيجاز إلى دور الذكاء الاصطناعي في حل هذه المشكلات.

ويتبع ذلك فصل نناقش فيه التعريب الآلي للأسماء الأجنبية: مشكلاته واقتراحات لبعض الحلول المجربة.

القسم الثالث: صناعة المعاجم العامة والمتخصصة، يتكون هذا القسم من ثلاثة فصول: صناعة المعجم العام، وأساليب التعامل مع المصطلحات العلمية، ثم نختم ببنوك المصطلحات.

القسم الرابع: يعرض للسانيات المدونات اللغوية والمدونات اللغوية: تعريف بها ومواصفاتها، وطرائق بنائها. ثم نستعرض بإيجاز أساليب الاستفادة منها في المجالات العلمية والعملية المختلفة. وقد ألحقنا بهذا الفصل ملحقان: قائمة بأهم المدونات اللغوية الخاصة باللغة العربية المعروفة، وآخر يقدم المصطلحات الرئيسة، نظرًا لحداثة هذا التخصص وقلة المراجع العربية فيه، وتيسرًا على القارئ العربي للدراسات باللغة الإنجليزية، إضافة إلى الراغبين في الكتابة في الموضوع أو ترجمة الدراسات الأجنبية إلى اللغة العربية. وقد جاء هذا الملحق في صورة معجم للمصطلحات الإنجليزية ومقابلاتها العربية، مع شرح وجيز باللغة العربية لكل من هذه المصطلحات. ويتبع ذلك مسرد ثنائي اللغة، مرتب حسب الكلمات العربية.

ونختم الكتاب بورقة تلخص أهم المجالات التي يمكن أن يخدم فيها الحاسوب اللغة العربية.

وكلي أمل أن يجد القارئ ولو بعض ما وجدته من متعة وفائدة في إعداد هذه الدراسات.

د. محمود إسماعيل صالح

(أستاذ اللسانيات التطبيقية)

الرياض

في شهر ذي القعدة، ١٤٤٢هـ (يونية، ٢٠٢١م)

مدخل الكتاب
الحاسوب واللغة

المقدمة:

من الأمور التي لا يجهلها أحد أن الحاسوب هو أداة القرن الحالي دون منازع. فقد دخل تقريبا كل بيت ومؤسسة عامة وخاصة، وأدى استخدامه في مجالات الحياة العلمية والعملية إلى تطورات كبيرة في هذه الحقول المختلفة. ومن هذه المجالات ميدان اللغة التي تميز بها الإنسان عن سائر الحيوانات. ويستفاد من الحاسوب في النشاط اللغوي بوجوه متعددة، يمكننا تلخيصها في ثلاث صور رئيسة هي: الاستعمال العام والاستعمال الخاص والبحث اللساني الحاسوبي المتخصص.

١- الاستعمال العام للحاسوب في النشاط اللغوي:

لعل من أهم مجالات الاستعمال العامة وأشيعها مايلي:

١-١- معالجة النصوص أو تنسيق الكلمات word processing:

ويقدم لها هذا النوع من البرمجيات إمكانات الكتابة وتوابعها مثل التدقيق الإملائي والتدقيق النحوي والمساهمة في دقة التعبير المعجمي من خلال معجم المترادفات أو المكنز thesaurus، إضافة إلى إمكانية إحصاء عدد ألفاظ النص وترتيبها sort (عند معاملتها على شكل فقرات) والبحث عنها في النص كذلك find (كما نجد في برنامج Word)، وغير ذلك من الإمكانات، مما يوفر على المستخدم الكثير من الجهد والوقت.

١-٢- عرض النصوص لأغراض الشرح والمناقشة الجماعية والتدريب (برنامج power point مثلا).

١-٣- استخدام الشابكة (الإنترنت) لأغراض مختلفة، مثل البحث عن المعلومات، والحصول على نسخ من بعض المطبوعات، والبحث المعجمي والمراجع اللغوية الأخرى، والوصول إلى المدونات المتاحة على الشابكة.

١-٤- التعليم والتقويم:

منذ العقد السابع من القرن الماضي بدأ استخدام الحاسوب في تعليم وتعلم اللغات، وهو ما يطلق عليه تعلم اللغة بمعاونة الحاسوب Computer Aided (Language Learning (CALL : ويشمل ذلك تعليم عناصر اللغة ومهاراتها بصور مختلفة ،كذلك تصميم الاختبارات اللغوية وتقديمها وتقويمها (انظرمحمود صالح: " تعليم اللغات باستخدام الحاسوب" في الملاحق).

١-٥- الترجمة:

ونجد هنا استخدامات عديدة للحاسوب، منها ما يلي: الترجمة الآلية أو الترجمة بمعاونة الحاسوب machine/ computer aided translation على الشابكة أو عن طريق برمجيات للاستعمال على الحاسوب الشخصي، وقواعد بيانات المصطلحات وبنوك المصطلحات، و"ذاكرات الترجمة translation memory" والمعاجم الإلكترونية (أحادية اللغة، ثنائية اللغة، المكانز thesauri– معاجم المترادفات والمعاني) (انظر: محمود صالح "الحاسوب في خدمة الترجمة: طرق الاستفادة والتقويم" و"بنوك المصطلحات والمعاجم الإلكترونية".)

١-٦- التعرف على النصوص المكتوبة:

ويشمل ذلك التعرف على المحارف البصرية optical character recognition OCR التي تحول النصوص المكتوبة (مطبوعة أو يدوية) إلى مادة رقمية ، ومن أهم أوجه الاستفادة من هذا النوع من البرمجيات تخزين النصوص المطبوعة في قواعد المعلومات والمدونات اللغوية.

١-٧- توليف الكلام:

ويسمى أحيانًا توليف الكلام speech synthesis لتحويل النص الرقمي إلى كلام منطوق.

ومن أمثلة الاستفادة من برامج قراءة النصوص المكتوبة شفويا ما يقدمه الحاسوب من خدمة للمكفوفين، حيث يغنيهم عن النصوص المكتوبة بنظام برايل Braille أو النصوص المسجلة صوتيا. ونجد فائدة ذلك أيضًا في الترجمة الشفوية.

١-٨- التعرف على النصوص الشفوية:

ويعتمد على تقنية التعرف على الكلام speech recognition، ثم(أ) تحويله إلى نص مكتوب في حالة الإملاء الآلي أو(ب) تنفيذ الأوامر المنطوقة، أو (جـ) "فهمه" في حالات الاستعلام الشفوي مثلا، وفي الترجمة الشفوية. كماويستخدم الأخير تقنية أخرى أيضا هي تقنية الذكاء الاصطناعي artificial intelligence (AI).

٢-البحث اللغوي:

هناك لا شك مجالات كثيرة يستخدم فيها الحاسوب في البحث اللغوي أوالبحث اللغوي النفسي. لعل من أبرزها:

٢-١- توليف الكلام والبحث اللغوي:

ويشمل ذلك استخدام توليف الكلام في الدراسات الصوتية وفي الدراسات اللغوية النفسية للتعرف على الملامح الصوتية التي يستفيد منها الإنسان في التعرف على الأصوات اللغوية، من ثم اللغة المنطوقة. ويستخدم الباحثون لذلك برنامج توليف الكلام speech synthesizer (Whitney, ١٩٩٨:١٤٨-١٥١)، وكذلك في عرض النصوص بسرعات مختلفة في دراسات التعرف على النصوص المكتوبة (انظر Whitney, ١٩٩٨; Harely, ٢٠٠٨)، وذلك في اللسانيات النفسية.

٢-٢- لسانيات المدونات اللغوية corpus linguistics:

لعل هذا الحقل هو أهم مجال في البحث اللغوي في العقود الأخيرة الماضية يقوم فيه الحاسوب بدورحيوي في تطوير البحث اللساني على شتى المستويات وفي شتى المجالات، كما سنرى في الفصل الخاص بالمدونات اللغوية..

٢-٣- قواعد البيانات والبحث اللغوي:

استخدامات برمجيات قواعد البيانات Data Base Management (Access):

يعتبر بعضهم المدونات اللغوية مثالا لاستعمال قواعد البيانات. ولكن هناك أمثلة أخرى لاستعمالها في البحث اللغوي، مثل تحليل الأخطاء، وذلك باستعمال قاعدة البيانات في تصنيف الأخطاء وفقا لأنواعها

وفئاتها ومستوى الدارسين وخلفياتهم والمهارات المختلفة وغير ذلك من معلومات. ولا شك أن هناك استعمالات أخرى كثيرة لقواعد البيانات في البحث اللغوي، إضافة إلى فوائدها في مراحل العمل المعجمي.
مثال لاستعمال قاعدة البيانات في تحليل الأخطاء:
ملف file: أخطاء الدارسين العجم
من سجلات records:
من حقول field:
حقل ١: العبارة الخاطئة (الكتاب جديدة)
حقل ٢: فئة الخطأ (نحو، إملاء، معجم)
حقل ٣: نوع الخطأ (حذف، إضافة، مطابقة...)
حقل ٤: مستوى الدارس (مبتدىء، متوسط، متقدم)

٣- اللسانيات الحاسوبية:

يخلط بعض الباحثين بين هذا المجال والمجال السابق، نظرا لأوجه التشابه والتداخل بينهما. ولكن يستحسن التفريق بين هذين الحقلين، فاستخدام الحاسوب في البحث اللغوي لايتطلب بالضرورة إلمام الباحث اللساني بالبرمجة الحاسوبية، فهو غالبا ما يسميه بعض الكتاب بالإنجليزية ordinary working linguist (OWL) أي الباحث اللغوي العادي.

أما اللسانيات الحاسوبية computational linguistics فتجمع بين اللسانيات وعلوم الحاسوب computer science (تخصص البرمجة الحاسوبية). وهوعلى الرغم من اعتباره تخصصا في اللسانيات أرى أنه أساسا نشاط حاسوبي يعتمد على المعلومات اللغوية. غير أن ذلك لا يمنع من وجود بعض اللسانيين المتمرسين في البرمجة الحاسوبية. ولا شك أن أي مشروع لساني حاسوبي ناجح يجب أن يجمع بين الخلفيتين الحاسوبية واللغوية، إما في الباحث الواحد أو بتعاون المتخصصين في كل من اللسانيات والحاسوب.

من المعروف أن المتخصصين في اللسانيات الحاسوبية هم الذين يصممون برمجيات ما يعرف بمعالجة اللغات الطبيعية natural language processing (NLP)، والتي تشمل: الترجمة الآلية وفهم النصوص والقراءة الآلية والتعرف على المحارف والأصوات وإنتاج الأصوات آليا والإملاء الآلي والتدقيق الإملائي

والنحوي والمكشاف السياقي concordancer وغيرها من البرمجيات مما يحتاج إليه الباحث في مجال لسانيات المدونات، مثل الوسم tagging والإعراب parsing الآليين، مما سنعرض له في ثنايا الأقسام التالية.

(لأمثلة من الدراسات اللسانية الحاسوبية، انظرمثلا: المحتسب وقهوجي، وحمادة، وراغين، والسلمان وآخرون في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية (٢٠٠٧).**

وجدير بالذكر أن برمجيات فهم النصوص أو الحوار مع الآلة تحتاج إلى تقنيات الذكاء الاصطناعي artificial intelligence (AI)، وهو تخصص أيضا يقع في مجال علوم الحاسوب. (انظرمثلا: "تطبيق نظام الخبرة في العلوم الشرعية باللغة العربية: نظام خبير في علم المواريث الإسلامية" في **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية (٢٠٠٧).**

٤-مجالات بحثية غير واضحة المعالم:

هناك دراسات يصعب تصنيفها ضمن الفئات السابقة، منها مثلا:

(أ) ما يقدمه بعض اللسانيين من دراسات بهدف خدمة البحث والنشاط اللساني الحاسوبي. نضرب على ذلك أمثلة: عملا محمود إسماعيل الصيني "نحو معجم عربي للتطبيقات الحاسوبية" و "المعجم ثنائي اللغة في الترجمة الآلية"، كذلك دراسة وفاء فايد (٢٠٠٧أ)، وكذلك نهاد الموسى (٢٠٠٠).

(ب) كذلك نجد صعوبة أحيانا في تصنيف المشروعات البحثية التي تنتقد أو تقوم المشروعات اللسانية الحاسوبية، كما في الدراسات التقويمية للمحللات الصرفية الآلية أو لمشروعات أو برمجيات الترجمة الآلية، مثل محمود صيني "الترجمة الآلية: إمكاناتها وحدودها"و وغيرها.

(جـ) وكذلك الدراسات التي يتعاون فيها اللسانيون والحاسوبيون، كل يدلي بدلوه في الدراسة بما يقع ضمن مجال تخصصه، كما نجد في دراسة عبدالعزيز والحملاوي الخاص بالمعجم المحوسب.

(د) أخيرا، يدخل في إطار مشكلة التصنيف الدراسات المسحية التي تحاول أن تجمع عددا من الدراسات المتعلقة باللغة والحاسوب، مثل داود عبده، أو تحاول حصر الأعمال البحثية في المجالين الثاني والثالث أعلاه، من أمثال عمل عبدالرحمن العارف. (انظر قائمة المراجع.)

وقد يفسر لنا كل ذلك مشكلة تصنيف النشاط البحثي فيما يتعلق بالدراسات التي تتضمن البحث اللساني والحاسوب بصور مختلفة. فلا نستطيع أن نلوم الباحثين الذين يجمعون كل هذه الضروب من الدراسات تحت مظلة "اللسانيات الحاسوبية".

وأرى أنه يمكننا تسمية هذا المجال باللسانيات الموجهة حاسوبيا، حيث إن الذين يقومون بمثل هذه الدراسات هم من اللسانين غالبا ممن لديهم اهتمام باستخدامات الحاسوب.

المراجع

أبو العزم، عبدالغني " اللغة العربية والمعالجة الآلية: برامج صخر نموذجا") في الموقع: www.aljabriabed.net/n٣١_٠٤abualazm.(٢).htm

أبو العزم، عبدالغني. الحاسوب والصناعة المعجماتية، **مجلة اللسان العربي**، العدد ٤٦، ١٩٩٨م، صص٢٨-٣٩.

بايزيد، ليلى (٢٠١٢) "التراص النصي في المقال الموسوم (فتاة مهمة) لسليمان العودة بين الشكل والمضمون" في: **دراسات في علم اللغة النصي: مقاربة تطبيقية على مدونة صحيفة الجزيرة** (٢٠١٢) الرياض: جامعة الأميرة نورة، ٢٩٥-٣٦٤.

بن يشو، جيلالي "حوسبة المعجم العربي: الواقع والآفاق" في: مجلة **أرتين – المرجع الأول لطلاب الأدب.** Art-En.com (تاريخ الوصول ٢٠١٣/٢/٢٢)

البواب، مروان (٢٠٠٩) أثر التقانات الحديثة في تجديد المعجم العربي (تاريخ الإضافة ٢٠٠٩/٣/١٨) www.aluka.net

البواب، مروان (٢٠١٢) محركات البحث في النصوص العربية. من بحوث المؤتمر السابع لمجمع اللغة العربية بدمشق (تاريخ الإضافة ٢٠١٢/٩/٢٣م) www.aluka.net/literature_language/٠/٧٧٢٨/

بوقرة، نعمان (٢٠١٢) "استراتيجيات الإقناع في الخطاب الصحفي السعودي- دراسة نصية تداولية" في: **دراسات في علم اللغة النصي: مقاربة تطبيقية على مدونة صحيفة الجزيرة** (٢٠١٢) الرياض: جامعة الأميرة نورة، ١٩-٩٠.

التعابير الاصطلاحية في ضوء النظريات اللسانية الحديثة: دراسة تطبيقية على مدونة صحيفة الجزيرة. الرياض: جامعة الأميرة نورة.

الثبيتي، عبدالمحسن عبيد (٢٠٠٧) "استخدام ذخائر النصوص لاستخلاص المصطلحات المتخصصة" في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق** البحثية، صص ٣١-٣٨.

الثنيان، نوال (٢٠١٢) "الإحالة الضميرية- دراسة نحوية نصية" في: **دراسات في علم اللغة النصي: مقاربة تطبيقية على مدونة صحيفة الجزيرة** (٢٠١٢) الرياض: جامعة الأميرة نورة، ١٩٣-٢٩٣.

الجلاب، محمد فتحي (١٤٣٢هـ) الاستخلاص الآلي للمحتوى العربي على شبكة الإنترنت بين الواقع والمأمول، في: **السجل العلمي لمؤتمر المحتوى العربي**، صص ٨٥٣-٨٩.

حاج صالح، عبدالرحمن (١٩٩٩) "ورقة حول مشروع الذخيرة اللغوية " في: **مجلة اللسان العربي**، العدد ٤٨.

الحلوة، نوال (٢٠١٢) " أثرالتكرار في التماسك النصي- مقاربة معجمية تطبيقية" في: **دراسات في علم اللغة النصي: مقاربة تطبيقية على مدونة صحيفة الجزيرة** (٢٠١٢) الرياض: جامعة الأميرة نورة، ٩١-١٩٠.

الحناش، محمد "برنامج لساني-حاسوبي للتعرف الآلي على التعابير المسكوكة في اللغة العربية" في: مجلة **التواصل اللساني**، ملحق سلسلة الندوات، المجلد ٣، سنة ١٩٩٦، ص ٨٩.

دراسات في علم اللغة النصي: مقاربة تطبيقية على مدونة صحيفة الجزيرة (٢٠١٢) الرياض: جامعة الأميرة نورة.

الربيعة، مها سليمان (١٤٣٣هـ) الدليل المرجعي للذخيرة النصية الفصحى لجامعة الملك سعود. *ksucorpus.ksu.edu.sa/ar/*

رشوان، محسن والمعتز بالله السعيد (تحرير، ٢٠١٩أ) **مقدمة في حوسَبة اللغة العربية**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

رشوان، محسن والمعتز بالله السعيد (تحرير، ٢٠١٩ب) **الموارد اللغوية الحاسوبية**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

الزركان، محمد علي (١٩٩٣) "اللسانيات و برمجة اللغة العربية في الحاسوب"، في: **وقائع المؤتمر الدولي الأول حول اللغة العربية والتقنيات المتقدمة**، الرياض، ١٠-١٤/٥/١٩٩٢، الرياض: مكتبة الملك عبدالعزيز، ١٩٩٣م، صص ٥٥-.

سالم، ثناء محمد (٢٠١٢) "الأبعاد التداولية للتعبيرالاصطلاحي" في: **التعابير الاصطلاحية في ضوء النظريات اللسانية الحديثة: دراسة تطبيقية على مدونة صحيفة الجزيرة**، صص ١٧-١٢٨.

السعيد، المعتزبالله (٢٠١٩تحرير). **العربية والذكاء الاصطناعي**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات المنعقدة في مكتبة الملك عبدالعزيز بالرياض (٨-١٢ /١١/ ١٤١٢هـ). الرياض: مكتبة الملك عبدالعزيز العامة.

السجل العلمي لندوة تقنية المعلومات والعلوم الشرعية والعربية المنعقدة في جامعة الإمام محمد بن سعود الاسلامية (١٦-١٤٢٨/٢/١٧هـ) الرياض: جامعة الإمام محمد بن سعود الإسلامية.

السجل العلمي: مؤتمر المحتوى العربي في الانترنت: التحديات والطموحات – المجلد الثاني (٥-١٤٣٢/١١/٧هـ) الرياض: جامعة الإمام محمد بن سعود الإسلامية،

السلمان، عبدالملك ومنصور الغامدي وحسن الصبي (٢٠٠٧) نظام حاسوبي لرومنة الأسماء العربية. في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية**، صص ٢١٥-٢٢٨.

صالح، محمود إسماعيل (٢٠٠٩) "الحاسوب في خدمة المترجمين"، دراسة مقدمة إلى: مؤتمر الترجمة والتعريب في المملكة العربية السعودية، الرياض (٢٨-٣٠ ديسمبر ٢٠٠٩).

الصيني، محمود إسماعيل (١٤١٢هـ). نحو معجم عربي للتطبيقات الحاسوبية. في: **السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات المنعقدة في مكتبة الملك عبدالعزز بالرياض (٨-١٢ /١١/ ١٤١٢هـ)، صص ٥١١-٥٢١.**

الصيني، محمود إسماعيل (١٩٨٩) "المعاجم في الترجمة الآلية" في: **اللسانيات العربية والإعلامية**. تونس:مركز الدراسات الاقتصادية والاجتماعية، ١٩٨٩م، ١٩٦.-١٨٣

الصيني، محمود إسماعيل (١٩٩١) الترجمة الآلية: إمكاناتها وحدودها " في: **المجلة العربية للثقافة** (تونس)، ع ٢١ (سبتمبر) ١٩٩١م، ١٤٤.-١٣٢

الصيني، محمود إسماعيل صالح (١٩٩٤) "بنوك المصطلحات الآلية والمعاجم الألكترونية": في**: السجل العلمي للندوة الثانية لتعريب الحاسوب (٣ -٢٧. مارس ١٩٩٤م)**، ٣٣٤.-٣١١

العارف، عبدالرحمن بن حسن "توظيف اللسانيات الحاسوبية في خدمة الدراسات اللغوية العربية: جهود ونتائج" في: **مجلة مجمع اللغة العربية الأردني**، العدد ٧٣.

عاليم، محمد (١٤٢٨هـ) المعجم العربي في ضوء اللسانيات الحاسوبية. في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية**، صص ١٥٧-١٦٦.

عبد العزيز، محمد حسن، محمد يونس الحملاوي والمعتز بالله السعيد طه (٢٠٠٨) المعجم الحاسوبي للغة العربية، بحث مقدم في: الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية المنعقد بمدينة الملك عبدالعزيز للعلوم والتقنية في الرياض (أبريل ٢٠٠٨).

عبده، داود عطية (١٣٩٩هـ) **المفردات الشائعة في اللغة العربية**، الرياض: جامعة الملك سعود،

عبده، داود عطية (١٤٣١هـ) **في اللغة والحاسوب: الترجمة وتدقيق الإملاء بين الإنسان والآلة**. عمان: دار جرير للنشر والتوزيع.

العريان، يوسف سالم (تحرير، ٢٠١٩) **تطبيقات الذكاء الاصطناعي في خدمة اللغة العربية**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

علي، نبيل (١٩٨٧) اللغة العربية والحاسوب: **مجلة عالم الفكر**، م ١٨، ع ٣، ص ٧٢ـــ

علي، نبيل (١٩٨٨) **اللغة العربية والحاسوب**. الكويت: مؤسسة تعريب.

علي، نبيل (١٩٩٠) نظم المعلومات: المشكلات والحلول. في: ندوة **"استخدام الحاسوب في العلوم الشرعية"**، صص ٢٠٧-٢٢٧.

علي، نبيل (١٩٩٤) العرب وعصر المعلومات، **سلسلة عالم المعرفة**، العدد ١٨٤، المجلس الوطني للثقافة والفنون والآداب، الكويت. (ينظر الفصل التاسع من الكتاب.)

عمر، أحمد مختار (١٩٨٨) **البحث اللغوي عند العرب مع دراسة في التأثير والتأثر**، ط٦. القاهرة: عالم الكتب.

العناتي، وليد أحمد "اللسانيات الحاسوبية العربية: رؤية ثقافية" **مجلة فكر ونقد** - المغرب. العدد ٨٢.

العناتي، وليد أحمد وخالد الجبر (٢٠٠٧) **دليل الباحث إلى اللسانيات الحاسوبية العربية**. عمان: دار جرير للنشر والتوزيع.

العصيمي، صالح بن فهد (تحرير، ٢٠١٥) **المدونات اللغوية العربية: بناؤها وطرائق الإفادة منها**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

غازي، عز الدين (٢٠٠٧) قواعد المعطيات المعرفية للمصطلحية العربية: مشروع مقترح. في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية**، صص ٢٣-٣٠.

الغامدي، عبدالله شرف و بدرية سليمان الفرهود (٢٠٠٧) أداة ويب معتمدة على عملية التحليل الهرمي للحصول على معجم عربي موحد لتقنية المعلومات. في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية**، صص ٣٩-٥٠.

الغامدي، منصور محمد "قاعدة بيانات الصوتيات العربية وقراءة الشفاه"
www٢٦.brinkster.com/mghamdi

فايد، وفاء كامل (٢٠٠٠) "بعض مظاهر تغير الصيغ الصرفية في العربية المعاصرة" في: **بحوث الندوة الدولية للمعاجم اللغوية العامة والمختصة**.(المنعقدة في الكويت في ١٤-١٧ مارس، ١٩٩٩)

فايد، وفاء كامل (٢٠٠٧أ) المتطلبات اللغوية لمعالجة التعابير الاصطلاحية العربية معالجة آلية، في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية**، صص ١١-٢٢.

فايد، وفاء كامل (٢٠٠٧ب) **معجم التعابير الاصطلاحية في العربية المعاصرة**. القاهرة: المؤلف.

الفلاج، نوال (٢٠١٢) "ظاهرة الاقتران الدلالي – دراسة معجمية تطبيقية" في: **التعابير الاصطلاحية في ضوء النظريات اللسانية الحديثة: دراسة تطبيقية على مدونة صحيفة الجزيرة**، صص ١٧٧-٢٧١.

الفيفي، عبدالله بن يحي (تحرير، ٢٠١٩) **خوارزميات الذكاء الاصطناعي في تحليل النص العربي**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

القحطاني، سعد بن هادي. "تحليل اللغة العربية بواسطة الحاسب الآلي" **مجلة علوم اللغة**، م ٥، ع ٣، ص ٢٢٨- (القاهرة: دار غريب.)

الكمار، رأفت (٢٠٠٧) **الحاسوب وميكنة اللغة العربية**. دار الكتب العلمية للنشر والتوزيع.

المالكي، رائدة وهند القحطاني (٢٠١٢) "العلاقات الدلالية للتعبيرات الاصطلاحية في مجال الرياضة" في: **التعابير الاصطلاحية في ضوء النظريات اللسانية الحديثة: دراسة تطبيقية على مدونة صحيفة الجزيرة**، صص ١٢٩-١٧٧.

مراياتي، محمد، يحي مير علم، محمد حسن طيان (١٩٩٦) **المعجم الحاسوبي: إحصاء الأفعال العربية في المعجم الحاسو**بي. بيروت: مكتبة لبنان.

مهديوي، عمر (١٤٣٢هـ) المقاربة الحاسوبية للصرف العربي: قراءة في الحصيلة والآفاق. في: **السجل العلمي لمؤتمر المحتوى العربي**، صص٩٩٩-١٠٢٧.

موسى، علي حلمي (١٩٧٢). **دراسة إحصائية لجذور لسان العرب باستخدام الكمبيوتر**. الكويت: جامعة الكويت.

موسى، علي حلمي (٢٠٠١) "حوسبة التراث العربي" محاضرة ألقيت في مجمع اللغة العربي الأردني في ١٧ نيسان (أبريل)، ٢٠٠١. /-٠٩-١٠-٠٢-٢٠٠٩/www.majma.org.jo/majma/index.php ٣٥-٢٨/٢٦٠-١٩-١.html

موسى، علي حلمي و عبدالصبور شاهين.(١٩٧٣) **دراسة إحصائية لجذور تاج العروس. باستخدام الكمبيوتر**. الكويت: جامعة الكويت.

موسى، علي حلمي (١٩٧١) **دراسة إحصائية لجذور معجم الصحاح باستخدام الكمبيوتر**. الكويت: جامعة الكويت.

موسى، علي حلمي (٢٠٠٧) **ألفاظ القرآن الكريم: دراسة علمية تكنولوجية**. القاهرة: الشركة المتحدة للطباعة والنشر والتوزيع.

الموسى، نهاد (٢٠٠٠) **العربية: نحو توصيف جديد في ضوء اللسانيات الحاسوبية**. بيروت: المؤسسة العربية للدراسات والنشر.

ندوة "استخدام الحاسوب في العلوم الشرعية" تحت إشراف مجمع الفقه الإسلامي (٢٤-٢٦ ربيع الآخر، ١٤١١هـ). جدة: البنك الإسلامي للتنمية.

الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية (٢٩/١٠-١١/٢/ ١٤٢٨هـ/ ١٠-٢٠٠٧/١١/١٢). الرياض: مدينة الملك عبدالعزيز للعلوم والتقنية.

النشرتي، مؤمن (١٤٣٢هـ) التحديات التي تواجه محركات البحث في استرجاع المحتوى العربي على الإنترنت: دراسة تحليلية، في: **السجل العلمي لمؤتمر المحتوى** العربي، صص ٧١٥-٨١٤.

هلال، يحي (١٩٩٠) الحاسوب في خدمة الحديث النبوي الشريف. في: ندوة **"استخدام الحاسوب في العلوم الشرعية"**، صص ٣٢٩-٣٤٢.

هليل، محمد حلمي (٢٠٠٨)"نحو معجم عربي معاصر" من بحوث الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية. الرياض: مدينة الملك عبدالعزيز للعلوم والتقنية.

هليل، محمد حلمي وسعد مصلوح وحشان العجمي (٢٠٠٠، تحرير) **بحوث الندوة الدولية للمعاجم اللغوية العامة والمختصة.(المنعقدة في الكويت في ١٤-١٧ مارس، ١٩٩٩)** الكويت: جامعة الكويت.

التعليم و التقويم

الحاسوب في تعليم اللغات

مقدمة:

من أهم المشكلات التي تواجه أنظمة التعليم في العالم عامة ندرة المعلم الجيد في حقول المعرفة المختلفة، وخصوصا في مجال تعليم اللغات الأجنبية. لذلك كان لابد من الاتجاه إلى التفكير في حل هذه المشكلة باستخدام وسائل التقنية الحديثة كالتعليم عن طريق التلفاز وباستخدام الحاسوب. ويسعدنا أن نقدم فيما يلي دراسة مختصرة لموضوع استخدام الحاسوب في تعليم اللغات، خاصة وأن الحاسوب قد بدأ يشق طريقه إلى سائر المؤسسات الحكومية والأهلية بل وإلى المنازل كذلك.

ويجب ألا يفوتنا أن نشير إلى الدور المتزايد الذي بدأت الشبكة العنكبوتية (الشابكة) تلعبها في سائر المجالات، بما في ذلك تعليم اللغة بوساطة الإنترنت Internet، وهو ما سنحاول التحدث عنه بإيجاز لاحقًا.

١- التعليم والحاسوب:

التعليم باستعمال الحاسوب غالبا ما يسمى بالتدريس بمساعدة الحاسوب Computer Aided/ Assisted Instruction (CAI). والمصطلح الأكثر تداولًا في مجال تعليم اللغات: تعلم اللغة بمساعدة الحاسوب Computer Assisted Language Learning (CALL) وهذا لا يعنى أنه لا يمكن الاعتماد على الحاسوب كلياً في تعليم بعض فروع المعرفة، ولكنه يعطى الصفة الأساسية لاستخدام الحاسوب في مجال التعليم: إنه وسيلة مساعدة. وإذا انطبق هذا الوضع على شيء، فأكثر ما ينطبق على استخدام الحاسوب في تعليم اللغات، حيث إن اللغة وسيلة تواصل بين مرسل ومستقبِل ومستقبِل ومرسل في نفس الوقت. ومشكلة الحاسوب أنه لا يستطيع أن يقدم ذلك البديل الكامل للمرسل والمستقبل في نفس الوقت، وذلك لأنه حتى في أفضل صورة لا يزال الحاسوب أداة مساعدة في تعليم اللغات.

والحاسوب أفضل وسيلة لإجراء التقويم والاختبارات الموضوعية، لأنه سريع ودقيق جدا، ويحلل النتائج في نفس الوقت مباشرة. وهذه هي الميزة التي تجعله مفيدا في التدريس كذلك، إذ من أهم مميزات الحاسوب في التدريس ما يسمى بالتغذية الراجعة feedback أ ما الميزة الثانية فهي فردية التعليم. ويتم ذلك من خلال ما يسمى بالتعليم المبرمج الخطي linear programmed instruction حيث يكون البرنامج بسيطًا، ولكن يستطيع أن يسير الدارس وفقًا لسرعته. وهناك أسلوب أكثر تطورًا، حيث يتيح لنا الحاسوب أن نعد برامج فردية الصبغة، وذلك ما يسمى بالتعليم المبرمج المتفرع . branched programmed instructionوفقًا لهذا

النوع من البرمجة يمكننا من خلال الحاسوب أن نقدم عشرات الدروس في نفس المادة الأساسية، تبعا لقدرات الطالب أو بالأحرى لاستجابة الطالب، فمثلًا إن أجاب الطالب الإجابة (أ) يأخذه الحاسوب إلى فقرة تختلف عن الطالب الذي تكون إجابته (ب) أو (جـ). فهو يسمح لنا بأن نصمم برنامجا مرنًا لطلاب من مستويات مختلفة ومن خلفيات مختلفة؛ بذلك يختلف خط سيركل طالب في البرنامج التعليمي نفسه وفقا لاستجاباته أو أدائه.

هذا وتشمل استخدامات الحاسوب في مجال تعليم اللغات عددًا من الحقول الأخرى، إضافة إلى التدريس والتدريب، مثل تقويم المواد التعليمية والتأليف وإعداد الاختبارات.

٢- تعليم اللغة:

٢-١- تعليم عناصر اللغة:

٢-١-١- تعليم قواعد اللغة وتراكيبها:

بقي هذا المجال لفترة طويلة واحدا من أشيع المجالات في برامج تعليم اللغات بمعاونة الحاسوب، حيث إن أكثر الموضوعات وضوحا وتحديدا وأقربها تناولا كما يبدو هو مجال التدريبات النحوية والصرفية. وهناك برامج كثيرة لتعلم القواعد والتراكيب. وهناك برامج بنيت على كتب معروفة مثل كتاب " مبادئ العربية المعاصرة " الواسع الانتشار في الجامعات الأمريكية، إضافة إلى برامج حاسوبية كثيرة متوافرة على الشابكة internet.

هذا ويمكننا تصميم التدريبات في أشكال مختلفة، منها برامج التدريب عن طريق:

(١)المزاوجة بين عمودين من أنصاف الجمل تكمل كل عبارة في العمود الثاني أخرى في العمود الأول مع التركيز على ظاهرة نحوية معينة، مثل كان وأخواتها أو إن وأخواتها. مثلًا نجد تحت العمود (أ) كان الجو، إن المكان... وفي العمود (ب) جميلُ اليوم، باردًا الآن... ويختار الطالب التكملة المناسبة (خبر مرفوع أو منصوب).

(٢)الاختيار من متعدد، حيث تعطى جملة ناقصة تتبعها عدة عبارات (أ ب ج) يختار الدارس واحدة منها لتكملة الجملة. مثلًا: يذهب التلميذ ---- المدرسة كل يوم. (أ) من (ب) في ج) إلى، وهكذا.

(٣)تكملة الجمل: يعطى الطالب جملة ناقصة ويطلب منه ملء الفراغ (استكمال الناقص) بناء على السياق المقدم في جملة سابقة.

مثال: أكمل الجملة التالية:

كان زيد فقيرا فلم يجد ما يأكله.

لوكان زيد غنيا....

(٤)التدريبات الآلية التقليدية مثل تدريبات التعويض (الاستبدال) والتحويل من مثبت مثلا إلى منفي أو من مبني للمعلوم إلى مبني للمجهول، والربط بين الجمل باستخدام أدوات العطف أو الشرط أواسم الموصول... إلى غير ذلك.

أ- التعويض: ضع الكلمة التي بين القوسين بدلًا مما تحته خط في الجمل التالية:

مثال: يكتب محمد الدرس. (يفهم) لتكون الإجابة: يفهم محمد الدرس.

يقرأ التلميذ الكتاب. (المعلِّم) لتكون الإجابة: يقرأ المعلمُ الكتاب.

ويمكننا أن نجعل التدريب أكثر صعوبة، بأن يقوم الدارس بإجراء تغيير بسيط على الجملة.

مثلًا: يكتب محمد الدرس. (فاطمة) لتكون الإجابة: تكتُب فاطمة الدرس، حيث يحول الفعل من صيغة المذكر إلى المؤنث.

ب- التحويل: من صيغة إلى أخرى:

- اجعل الجمل التالية منفية، كما في المثال:

المثال: سمع الطالب التعليمات. لم يسمع الطالب التعليمات.

كتب زيد الواجب....

- اجعل الجمل التالية مبنية للمجهول:

مثال: فهم التلاميذ الدرس. فُهِم الدرسُ.

سمِع المصلون الإقامة. سُمِعت الإقامة.

جـ- ربط الجمل: استعمال أداة العطف أو اسم الموصول مثلًا:

- اربط الجمل مستعملًا حرف العطف (و).

مثال: دخل المعلم الصف. جلس المعلم على الكرسي. وتكون الإجابة: دخل المعلم الصف وجلس على الكرسي.

- اربط الجمل مستعملًا اسم الموصول (الذي أو التي):

مثال: قرأت كتاب النحو. استعرت كتاب النحو من المكتبة. وتكون الإجابة: قرأت كتاب النحو **الذي استعرته** من المكتبة.

د- التوسعة: إضافة كلمة أو كلمات إلى الجملة.

ضع الكلمة التي بين القوسين في مكانها الصحيح في الجملة، مع تغيير ما يلزم.

استمعت إلى تلاوة. (جميل). وتكون الإجابة: استمعت إلى تلاوة **جميلة**.

(٥)الألعاب مثل لعبة التخمين، للتدريب على أدوات الاستفهام (من، ماذا، كيف، لماذا، أين...الخ)

وهناك برنامج مشهور للتدريب على صيغ التفضيل (أكبر، أصغر، أدق) وهو على شكل ذكر أوصاف ملامح مجرم مطلوب القبض عليه. مثلا يجد الدارس على الشاشة أمامه صورة وجه للحظات ثم يختفي ذلك الوجه. ويحاول الدارس أن يعيد رسم ذلك الوجه بتحديد ملامحه من قائمة من الكلمات (مثلا: أنفه أكبر، حاجباه أدق...) ويقوم الحاسوب بتنفيذ الوصف على الصورة الجديدة إلى تمكين الدارس من تعديلها حتى

تتطابق مع الصورة الأصلية (Jones and Fortecue ١٧-١٨)

يجدر بالذكر أن هناك برامج حاسوبية لتشكيل النصوص العربية، بدرجات متفاوتة من الدقة. ومن أشهرها برنامج "التشكيل الآلي" من شركة صخر، tashkeel.alsharekh.org.

وبرنامج "مشكال للنصوص العربية" tahads.com/mishkal

٢-١-١-٢ المراجع النظرية والتعليمية الأخرى:

تتيح لنا الشابكة الوصول إلى عشرات المواقع العربية التي تعالج الجوانب المختلفة من النحو العربي التي يمكن لكل من معلمي العربية ودارسيها الرجوع إليها في أي زمان ومكان. ومن هذه المواقع: شَبَكَةُ الفَصِيحِ لِعُلُومِ اللُّغَةِ العَرَبِيّةِ *www.alfaseeh.com/*

٢-١-٢-١- تدريس المفردات:

يبدو أن مجال المفردات خصب جدا للألعاب اللغوية الطريفة، إضافة إلى التدريبات التقليدية مثل ذكر المرادف والمضاد وملء الفراغات بالكلمات المناسبة أو المزاوجة بين المرادفات والأضداد وبين التعريفات والكلمات المناسبة لها.

- من أمثلة الألعاب لعبة "الكلمة الدوارة ".Wordspinيشاهد الدارس حروف كلمة ما في ترتيبات مختلفة بصورة متتابعة. وعلى الدارس أن يضغط على إزرا ر معين ليوقف هذا التتابع عند مشاهدته الكلمة في صورتها الصحيحة.

-وفي لعبة أخرى يختار الدارس مجالا معينا (مثل الخضروات) فتظهر كلمة على الشاشة من ذلك الحقل وتختفي. ثم يطلب من الدارس إعادة كتابة تلك الكلمة. ويستطيع الدارس أن يتحكم في سرعة البرنامج لمزيد من التحدي والاثارة.

ومن التدريبات والألعاب الأخرى:

- لعبة الكلمات المتقاطعة، والكلمات المخفية، حيث يؤلف الدارس من الحروف الموجودة في داخل مربع ما أكبر عدد من الكلمات مرتبة من أعلى إلى أسفل أو من اليمين إلى اليسار أو في أى ترتيب آخر.

- تركيب الكلمات من أجزاء متفرقة، حيث يحاول الدارس تركيب أكبر عدد ممكن هذه الأجزاء. مثلا: يجد الدارس (- ون، - ات، - اء) ويجد أيضا (طالب، صحر، مهندس، خضر) ليكون منها: طالبات، مهندسون، صحراء، خضراء) مثلا.

- تصنيف الكلمات: يختار الدارس مثلا عدة أصناف من الكلمات (زهور، أطعمه، فواكه... أو أسماء، أفعال...) فتظهر على الشاشة خانات مرقمة. ثم تظهر كلمات مختلفة (مثل: موز، ورد، أرز أو كتاب، قرأ)، على الدارس أن يضع كل منها في الخانة المناسبة من حيث صنفها.

- تدريب المشتركات اللفظية: يقدَّم للدارس كلمة في سياق معيّن ويطلب منه التعرف على المعنى المناسب للسياق.

مثلًا: "قرأت الجملة عينها". معنى "عين": نفس، حاسة النظر، بئر ماء.

تدريب المتلازمات اللفظية: نورد للدارس قائمة (أ) بكلمات وفي قائمة (ب) كلمات أخرى تلازمها أو تصاحبها، مثل:

(أ)	(ب)
جماعة	غنم
فريق	مصلين
قطيع	لاعبين

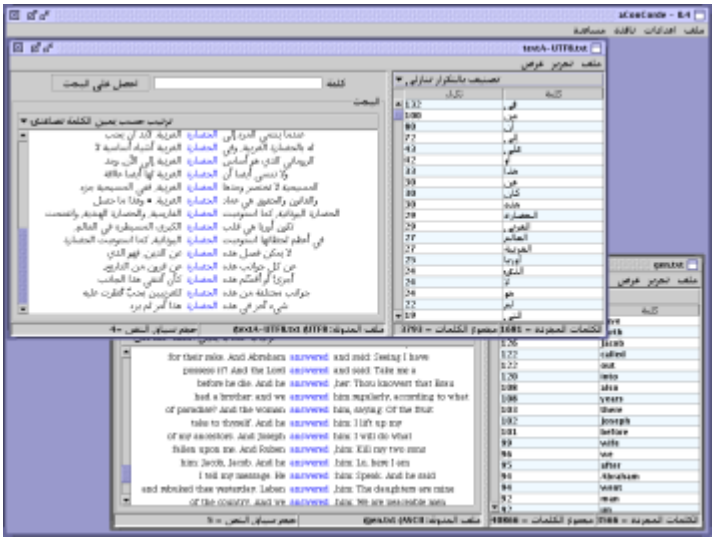
ويقوم الدارس بمواءمة كل كلمة من القائمة (أ) بما يناسبها من القائمة (ب) عن طريق القص واللصق(فيقول مثلًا: قطيع غنم.)

- الاشتقاق: وهو من خصائص العربية، حيث يطلَب من الدارس الإتيان بعدد معين (مثلًا ثلاث أو خمس كلمات) من جذر ما (ك ت ب) فيقول: (كاتب، مكتوب، كتابة).

- المواءمة بين التعبيرات الاصطلاحية ومقابلاتها (مثل "انتقل إلى الرفيق الأعلى" ومقابلها "توفي".)

- الترجمة: هناك برامج تستخدم الترجمة وسيلة للتدريب على المفردات وتنمية الثروة اللفظية. وهو ما نلاحظه في كثير من البرامج الوارد ذكرها في الملحق. (انظر الملحق.)

من البرمجيات التي شاع استخدامها مؤخرًا في تعليم المفردات (معانيها واستعمالاتها ومصاحباتها اللفظية) المكشاف السياقي concordancer، وهناك عدد من المكشافات يمكن استخدامها مع النصوص العربية، مثل aconcorde وغواص Ghawas. وأسلوب عملها هو تخزين النص (ما يسمى بالمدونة corpus)، ثم نطلب من البرنامج تحليل النص ليورد لنا جميع كلمات النص وتكراراتها. ويمكننا أن نطلب أن يورد كل كلمة في سياقاتها المختلفة، حيث نستطيع استخلاص معاني الكلمة والكلمات المصاحبة لها (أي الكلمات التي تسبقها والتي تأتي بعدها). ويمكن للمعلم أن يكلف الطلاب خصوصًا في المستويات المتقدمة أن يقوموا بالبحث بأنفسهم للحصول على المعلومات المطلوبة. انظر صورة الشاشة أدناه:



نرى في الشاشة الأولى جزءًا من قائمة الألفاظ في النص مع بيان تكرار كل منها. وفي الجانب الأيسر كشافًا سياقيًّا لكلمة "الحضارة" مع عدد من الكلمات قبلها وعدد منها بعد الكلمة "المفتاحية key word" الظاهرة باللون الأزرق. وفي الشاشة الثانية الشيء نفسه لنص إنجليزي.

(انظرصالح (٢٠١٥) "المدونات اللغوية وطرق الإفادة منها".)

٢-١-٢-٢- المعاجم الإلكترونية:

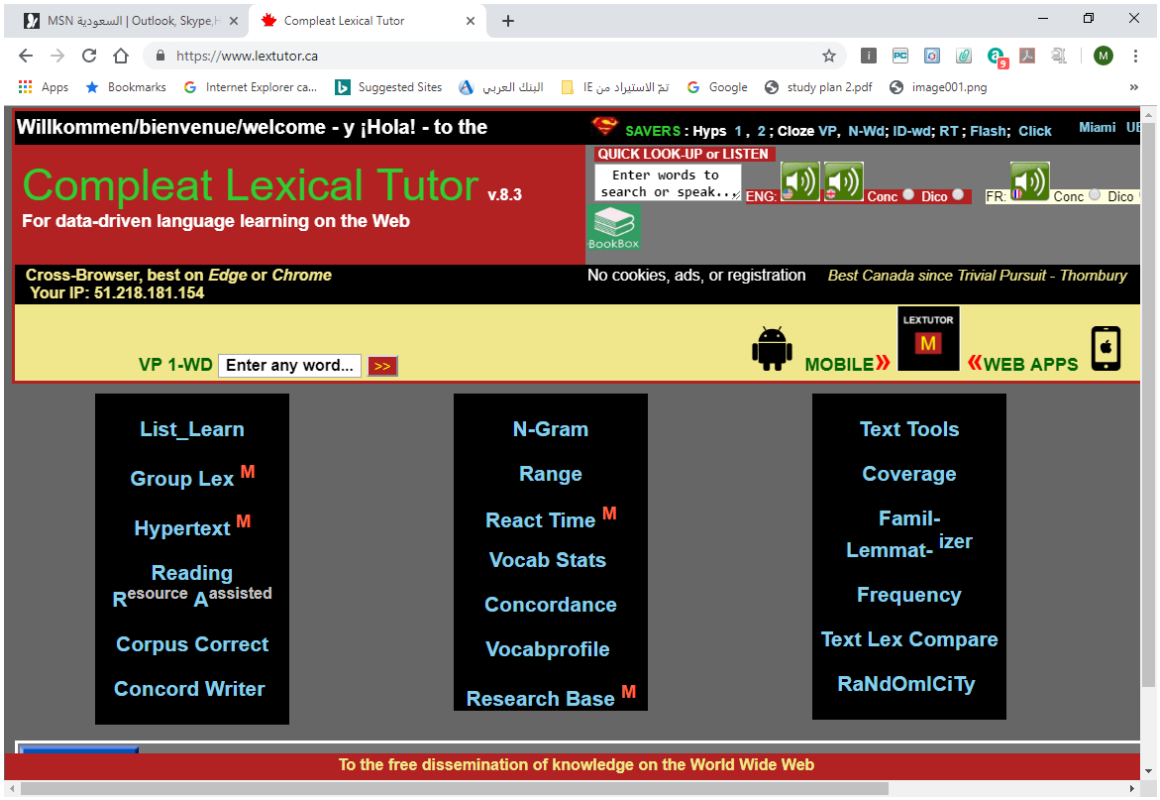
من الأدوات المفيدة جدًا في تعلم المفردات وتعليمها المعاجم الإلكترونية المختلفة، أحادية اللغة وثنائيتها. وهي متاحة في صورة أقرص مضغوطة وفي شكل معاجم متاحة على الشابكة. فمن حيث المعاجم أحادية اللغة معاجم الألفاظ المعروفة، مثل لسان العرب أو القاموس المحيط أو المنجد وغيرها، إضافة إلى المعاجم التفاعلية على الشابكة. ومن المعاجم ثنائية اللغة المورد ومعجم اللغة العربية المعاصرة، إضافة إلى عشرات المعاجم المتخصصة في مجالات مختلفة، مثل المعجم الطبي الموحد.

وتمتاز المعاجم الإلكترونية بسهولة البحث فيها، خاصة وأن بعضها يتيح البحث عن طريق الجذر (ك ت ب) أو الكلمة (مكتوب)، مما لا يتطلب معرفة الباحث بجذر الكلمة للبحث عنها. ومن المعاجم المفيدة جدًا " قاموس ومعجم المعاني متعدد اللغات والمجالات" https://www.almaany.com/

ومن أحدث المعاجم العربية المعجم المعاصر من إنتاج شركة صخر، lexicon.alsharekh.com

ثم هناك معاجم المترادفات والمعاني، ما يسمى بالإنجليزية thesaurus. وهي مفيدة في التأليف والترجمة، حيث تتيح للباحث التعرف على الكلمات المناسبة للتعبير عن أفكاره.

ومن البرامج المفيدة في تدريس المفردات في اللغة الإنجليزية حزمة compleat lexical tutor الذي تشتمل على عدد من الملفات المختلفة للتعليم والتدريب والاختبار. انظر www.lextutor.ca



٢-١-٣-تدريس الأصوات والنطق:

هناك أساليب كثيرة للاستفادة من الحاسوب في التدريب على الأصوات تعرفًا ونطقًا، بما في ذلك الشرح النظري في وصف الأصوات ومخارجها وتعريف بالرموز الصوتية (مثل الألفبائة الدولية IPA). ولعل أبسط مثال هو استخدام الأقراص المضغوطة CD التي تقدم النماذج المسموعة للدارس بغرض التدرب على الاستماع إلى النصوص اللغوية المختلفة ومحاكاتها في النطق.

هناك مثلًا برامج للتدريب على نطق الأصوات والكلمات والجمل يقوم الحاسوب فيها بعملية تقويم أداء الدارس كذلك. ومن أمثلة ذلك برنامج الكلام المرئي ((visible speech وهو برنامج يقدم للدارس جملا وعبارات منطوقة يحاول الدارس أن يحاكيها فيجد على الشاشة مقارنة بين أدائه وأداء النموذج. وهذا البرنامج يستخدم للتدريب على تنغيم الجمل، وليس النطق الصحيح للأصوات القطعية (الصوامت والصوائت). كذلك نجد أن بعض البرامج المتوافرة في الأسواق لتعليم اللغات المختلفة يصاحبها برنامج التدريب على النطق الصحيح تستخدم ما أشرنا إليه أعلام بإمكانية التعرف على الكلام speech

recognition، حيث يستمع الدارس إلى العبارات المختلفة ويحاكيها، فيقوم الحاسوب بتقويم أدائه: هل هو جيد أو سيئ أو مقبول، وهكذا.

من مصادر تعليم أصوات العربية ونطقها التي يغفل عنها كثير من الناس برامج تعليم تلاوة القرآن الكريم (إضافة إلى تسجيلات القرآن الكريم المعتادة)، وهي متاحة بوفرة على الشابكة وفي صور الأقراص المضغوطة CD. ومما لا شك فيه أن تعلم التلاوة خير وسيلة للتعلم الصحيح لنطق العربية وأصواتها. ومن البرامج المعروفة لتعليم تجويد القرآن الكريم برنامج "حفص"www.hafss.net وهو متاح للتحميل في صورة قرص مضغوط أو مدمج، كما يسمى أحيانًا.

٢-١-٤- تعليم قراءة الحروف وكتابتها:

منذ أكثر من ثلاثين عامًا طورت الباحثة فكتورين عبود من جامعة تكساس في أوستن برنامجًا لتعليم قراءة الحروف العربية وكتابتها. واليوم لازالت الجامعة المذكورة تقدم برنامجًا على الشابكة (الإنترنت) لتعليم الحروف إلى جانب عناصر اللغة العربية ومهاراتها المختلفة، كما هو وارد في الملحق في هذه الورقة. وهناك حاليًا عدد من البرامج المختلفة لهذا الغرض.

هنا يجب التنبيه إلى ضرورة التفريق بين الكتابة اليدوية واستعمال لوحة مفاتيح الحاسوب. فتعليم الكتابة اليدوية يحتاج إلى إمكانات خاصة مثل اللوحة التي نكتب عليها يدويًّا بقلم خاص لتتحول الكتابة إلى شاشة الحاسوب. كذلك إذا أردنا للحاسوب أن يقيم أداء الطالب فلا بد من وجود برنامج للتعرف على الحروف character recognition حتى يقارن بين ما كتبه الطالب وبين ما يجب أن يكون عليه شكل الحرف أو الكلمة.

٢-٢- تعليم مهارات اللغة

٢-٢-١- تعليم الاستماع:

٢-٢-١-١- المهارات الجزئية:

- تمييز الاصوات: في هذا البرنامج يستمع الدارس الى ثنائيات صغرى minimal pairs ويحدد ما إذا كانت

الكلمات التي يسمعها متماثلة أو مختلفة، موضحا ذلك عن طريق لوحة المفاتيح على الحاسوب. ويتم حساب إجابات الطالب الصحيحة والخاطئة.

- تحديد نوع التنغيم:

يستمع الدارس الى جمل مسجلة بأنواع من التنغيم مختلفة، ويحدد نوعها (مثلا: تدل على الإخبار أو الاستفهام أو التعجب...) ويدخل الإجابة في الحاسوب عن طريق لوحة المفاتيح، ليحدد الحاسوب صواب الدارس من خطئه ويقيم أداءه في ضوء ذلك.

- الإملاء:

يستمع الدارس الى نص أو مجموعة من الكلمات ويحاول كتابته عن طريق لوحة المفاتيح. ثم يقدم له الحاسوب النص بصورته الصحيحة. ويتم تقويم أداء الطالب (من حيث نسبة أخطائه وتقدمه مثلا).

٢-٢-١-٢- فهم المسموع:

من البرامج المعروفة برنامج إملاء الصور picture dictation: وهو يقدم للدارس على الشاشة مخططا لغرفة نوم مثلا به نافذة وباب وحروف تمثل مواقع قطع الأثاث المختلفة. يستمع الدارس إلى التسجيل الذي يحدد له مواقع قطع الاثاث المختلفة ثم تتلو ذلك مجموعة من الأسئلة عن موقع كل قطعة من الأثاث ليختار الدارس على لوحة المفاتيح الحرف المناسب للموقع... وهكذا.

وهناك برنامج (فهم الرسالة) حيث يستمع الدارس إلى رسائل هاتفية ثم يقوم بتنفيذ ماتطلبه منه الرسالة أو يجيب عن أسئلة تتعلق بما ورد فيها. (انظر Jones and Fortescue, ٨٠-٨١).

من الجدير بالذكر أن الشابكة (الإنترنت) من أفضل المصادر للتدريب على فهم المسموع، حيث تتيح لنا مشاهدة وسماع عشرات المحطات الإذاعية والتلفازية، إضافة إلى اليوتيوب You tube، في شتى اللغات وجميع الموضوعات.

٢-٢-٢- تعليم التعبير الشفوي:

٢

٢-٢-٢-١- المهارات الجزئية:

كما ذكرنا أعلاه لايقوم الحاسوب عادة إلا بتوفير المناخ اللازم أو المشجع على الكلام والحوار في معظم برامج التدريب على الكلام الشمولية. ولكن هناك برامج للتدريب على بعض المهارات الجزئية يقوم الحاسوب فيها بعملية التقويم كذلك. ومن أمثلة ذلك برنامج الكلام المرئي (visible speech) وهو برنامج يقدم للدارس جملا وعبارات منطوقة يحاول الدارس أن يحاكيها فيجد على الشاشة مقارنة بين أدائه وأداء النموذج. وهذا البرنامج يستخدم للتدريب على تنغيم الجمل، وليس النطق الصحيح للأصوات القطعية (الصوامت والصوائت). لكن هناك أيضًا برامج التدريب على النطق الصحيح تستخدم ما أشرنا إليه أعلاه بإمكانية التعرف على الكلام.

-٢-٢-٢-٢-المهارات الكلية:

يتمثل دور الحاسوب في التشجيع على التدريب الشفوي أو تهيئة الظروف التي تؤدي إلى الاستخدام الشفوي للغة من حوار ونقاش وجدال إلى غير ذلك. ولا يستطيع الحاسوب التقليدي تقويم عمل الطلاب ما لم يتم تحويل كلامهم إلى مادة مكتوبة، كأن يقدم الحاسوب مشكلة ما يتناقش فيها أفراد مجموعة من الدارسين ثم يكتبون الاجابة عن طريق لوحة المفاتيح في الحاسوب. ولكن تقويم ذلك لا يعتبر تقويما لأدائهم الشفوي.

غير أن هناك عددا متزايدا من برامج تعليم اللغة بمساعدة الحاسوب بدأت في الاستفادة من إمكانات مايسمى بالتعرف على الكلام speech recognition حيث يستطيع الحاسوب القيام بمقارنة أداء نطق الدارس بالنموذج المخزون في ذاكرته، ويتم بذلك تقويم أداء الدارس الشفوي.

وهناك عدد من البرامج التي تقدم مثل هذه المشكلات بالصورة والصوت أحيانا لتحفز الدارسين على التحدث أو تكملة الحوارات الناقصة أو لتمثيل الأدوار.

ونجد في البرامج المتطورة استخداما للحاسوب على الأقراص المرئية(الفيديوديسك) أو على الشابكة لتقديم صور حقيقية ومتحركة للمواقف المختلفة، مثل منظر سائح يتجول في سوق بلد أجنبي ويدور حوار بينه وبين الباعة أو بينه وبين المرشد وهكذا. ويقوم الحاسوب بتمثيل دور أحد هؤلاء المشاركين في الحوار، بينما يقوم الدارس بأداء دور الطرف الآخر.

٢-٢-٣- تدريس القراءة

٢-٢-٣-١- فهم المقروء

لابد لنا أن نذكر أن القراءة من المهارات التي يتم التدريب عليها بصورة شبه مستمرة في دروس اللغة بمعاونة الحاسوب، حيث إن المتَبع عادة هو أن معظم التعليمات والتدريبات بأشكالها المختلفة تتم غالبا عن طريق المادة المكتوبة على الشاشة. وقد سمى بعضهم ذلك التدريب بالقراءة العابرة (أي غير المقصودة) incidental reading. لكن هناك برامج تصمم خصوصا للتدريب على القراءة بجانبيها الآلي (التعرف على الحروف والرموز المكتوبة) والمعرفي (أي فهم المقروء). ويقسم جونز وفورتسكيو Jones and Fortescue تدريس القراءة إلى ثلاثة أنواع رئيسة:

(١)القراءة العابرة

(٢)فهم المقروء

(٣)التصرف في النصوص .text manipulation

كما يذكران أن هناك بعض التدريبات التى يمكن تسميتها بتدريبات ما قبل القراءة (أى قبل قراءة النص)، وهي تدريبات تمهد لقراءة النص، عن طريق أسئلة يحاول الدراسون الإجابة عنها قبل القراءة. ثم تتاح لهم فرصة الاطلاع على النص للتثبت من صحة إجاباتهم. بعد ذلك يقدم الحاسوب الإجابة الأولية للطلاب عن كل سؤال ثم الإجابة في صورتها الأخيرة تتلوها الإجابة الصحيحة. بذلك يتهيأ الدارس لقراءة النص والتدرب على فهمه كاملا.

والصورة التقليدية المعروفة هي طبعا عرض النص المقروء على الشاشة ثم إتباعه بالأسئلة فى صورها المختلفة:

- صحيح وخطأ، حيث يقرآ الدارس الجملة ويشير إلى كونها صحيحة أم خاطئة في ضوء النص. ويقترح أن نضيف احتمالًا ثالثا، وهو "لا ندري"، حيث يورد المعلم عبارة أو جملة لا يعرف صحتها أو خطؤها في ضوء النص المقروء.
- الاختيار من متعدد، حيث يورد المعلم عددًا من العبارات أو الجمل ويطلب من الدارس أن يختارها الصواب أو الأقرب إلى الصواب منها.

- ملء الفراغات، حيث تقدم جملا ينقص من كل منها كلمة أو عبارة يأتي بها الدارس، بناء على النص.
- تكملة الجمل والعبارات الناقصة، حيث يقرأ الدارس جملة مثلًا ناقصة ويأتي بالعبارة المناسبة لاستكمال الناقص فيها.

ومن التدريبات المستخدمة للتدريب على الفهم :

- إعادة بناء الجمل مختلفة الأجزاء، حيث يقدم الحاسوب أجزاء الجملة بدون ترتيب ومعها (قبلها و / أو بعدها) جمل توضح السياق. ثم يطلب من الدارس إعادة ترتيب تلك الاجزاء من نص أو حوار.

- ترتيب الفقرات المبعثرة لتكوين نص متماسك.

- اختيار العنوان المناسب للنص.

- اختبارات التتمة cloze test، حيث يورد الحاسوب نصا حذفت منه بعض كلماته وعلى الدارس أن يأتى بها. هذا ويمكننا أن نقدم للدارس قائمة من الكلمات يختار منها. وإذا كان النص طويلًا، فيجب أن نقسمه إلى فقرات، لكل فقرة قائمة من الكلمات يختار منها الدارس لإكمال الناقص.

٢-٢-٣-٢- التدريب على القراءة السريعة:

ويتمثل ذلك في تقديم الحاسوب لنص ما وفق سرعات مختلفة يختارها الدارس. يتلو النص أسئلة للفهم. وعن طريق زيادة سرعة عرض النص يتم تدريب القارئ على الإسراع في القراءة لملاحقة عرض النص على شاشة الحاسوب.

وهناك برامج تقدم الأسئلة قبل قراءة النص حتى يوجه القارئ اهتمامه على البحث عن إجابات لتلك الأسئلة.

٢-٢-٤- تعليم الكتابة

٢-٢-٤-١- المهارات الجزئية

لابد لنا هنا أن نفرق بين (١) تدريبات الكتابة الجزئية، مثل التعبير المقيد بأشكاله المختلفة والتدريبات النحوية وتدريبات الاستيعاب. وقد أشرنا إلى كثير منها في الأجزاء الخاصة بالتراكيب والنحو والقراءة. (٢) تدريبات الكتابة الشاملة.

سنبدأ أولا بالتدريبات الجزئية أو مايسمى أحيانا بتدريبات المهارات الفرعية subskills التى تساعد على التدريب على الكتابة أو التعبير التحريري.

أولا: التدريبات الجزئية:

١- الأسئلة والأجوبة:

وتشمل هذه التمارين التى تطلب من الدارس وضع أو توجيه أسئلة، تبعا للتعليمات.

مثال: ضع سؤالا عن العبارة التى تحتها خط في كل جملة مما يأتي:

-هاجر النبي صلى الله عليه وسلم من مكة الى يثرب.

-كانت الهجرة النبوية في السنة الثالثة عشرة من البعثة.

-بقي النبي صلى الله عليه وسلم وصاحبه في الغار ثلاثة أيام.

ثم هناك الأسئلة التي تتلو النصوص المقروءة والمسموعة، وتبدأ بأدوات الاستفهام المعروفة (من، أين، كيف، متى، لماذا...). ويطلب من الدارس الاجابة عنها.

٢-تكملة الجمل الناقصة:

تعطى للدارس مجموعة جمل ناقصة، أي تم حذف جزء من كل منها، ويطلب منه استكمالها. وقد تكون هناك سياقات محددة (جملة تامة قبل الجملة الناقصة مثلا).

مثال: أكمل الجمل الناقصة فيما يأتي:

- في رمضان من كل عام ـــــ من الفجر إلى غروب الشمس.

٣- المزاوجة بين أجزاء الجمل، حيث يكتب في أحد العمودين الأجزاء الأولى من كل جملة وفي العمود الثاني الأجزاء الأخيرة لتلك الجمل، بدون ترتيب. ويقوم الدارس بالمزاوجة بين الأجزاء المتلائمة تركيبيا ودلاليا، ثم يقوم بكتابة الجمل التامة.

٤- ترتيب الكلمات المبعثرة في جمل:

تقدم للدارس مجموعات من الكلمات. ويطلب منه أن يرتب كل مجموعة لتكوين جملة صحيحة. وقد يطلب

منه أحيانا إعادة صياغة بعض الكلمات (مثلا، تحويل صيغة الفعل ليطابق فاعله من حيث التأنيث والعدد، أو لمراعاة قواعد النصب والجزم).

٥- ترتيب الجمل لتكوين فقرات:

للتدريب على التنظيم المنطقي في التعبير، يعطى الطالب مجموعة من الجمل المترابطة بعد تغيير ترتيبها. ويقوم الدارس بترتيب الجمل في مواقعها الصحيحة، إما بعادة كتابتها مرتبة، أو بترتيب أرقامها (اختصارا للجهد والوقت).

٦- وضع علامات الترقيم المحذوفة:

يقدم للدارس نص حذفت منه علامات الترقيم. ويطلب منه أن يضعها، بناء على قواعد الترقيم وعلى المؤشرات cuesالموجودة في النص. مثلا في اللغة الانجليزية، إذا وجدنا حرفا كبيرا لغير أسماء الأعلام عرفنا أن ذلك دليل على بداية جملة ما. وفي سائر اللغات تدل أدوات الاستفهام على وجوب وضع علامة استفهام في نهاية الجملة... وهكذا.

ثانيا: الكتابة الشاملة:

من أهم الاستخدامات الحاسوبية في الكتابة، ما يعرف بمعالج الكلمات word processorأو منسق النصوص كما يسمى بالعربية أحيانا.وهو برنامج يجعل الحاسوب مثل الراقنة (الآلة الكاتبة)، لكن بميزات إضافية كثيرة أهمها ما يلي:

- إمكانات الحذف والإضافة والتعديل الفورية.

- إمكانية التقديم والتأخير لأي جزء من النص، وهو ما يعرف بعمليتي (القص) و(اللصق).

- إمكانية البحث عن أجزاء من النص والتعديل عليها بصورة فردية أو بصورة شموليةglobal . مثلا، إذا قرر الكاتب تغيير عبارة أو كلمة وردت في النص ولا يذكر موقعها، كل ما عليه هو إعطاء التعليمات للحاسوب للبحث عن تلك العبارة أو الكلمة. فيقوم الحاسوب باستعراض النص ويقف عند كل حالة ترد فيها العبارة المطلوبة ليقوم الكاتب بإجراء التعديلات عليها، بل ويمكنه أن يوفر الوقت والجهد باستخدام أمر (البحث مع التغيير لحالة واحدة أو للكل)، حيث يكتب العبارة المطلوب تغييرها، ثم العبارة بصورتها الجديدة فيقوم الحاسوب بالبحث والتغيير تلقائيا للعبارة أينما وردت في النص.

- التدقيق الاملائي:

تتوافر مع معظم برامج معالجة الكلمات برامج للتدقيق الإملائي spell checker، حيث يقوم الحاسوب بناء

على طلب من الكاتب بتدقيق النص إملائيا، والتنبيه إلى الكلمات التي يشك في صحة كتابتها، مع اقتراح كلمة أو كلمات يختار منها الكاتب في معظم الأحيان.

- المراجعة النحوية:

هناك برامج معالجة الكلمات في بعض اللغات مثل الانجليزية تقوم بالمراجعة النحوية، حيث تنبه إلى بعض الأخطاء المحتملة نحويا وأسلوبيا.

- معجم المترادفات أو المكنز thesaurus

يقدم العديد من برامج تنسيق النصوص الأجنبية إمكانية البحث عن مرادفات للكلمات، ليختار منها الكاتب الأوفق والأنسب للسياق.

- توفير قاموس ثنائي اللغة:

هناك برامج في العالم العربي تقدم للكاتب فرصة البحث عن المقابلات الأجنبية لعدد كبير من الكلمات العربية.

- المزج بين اللغات:

تسمح جميع برامج معالجة الكلمات ثنائية اللغة بمزج الكلمات والعبارات من اللغتين كالعربية وأخرى أجنبية (كالفرنسية أو الإنجليزية) جنبا إلى جنب.

- إمكانيات اختيار الأبناط المختلفة:

تسمح معظم برامج معالجة الكلمات باختيار الأبناط، بل والاحجام، المختلفة للنص ولأجزاء منه (مثل العناوين الرئيسية والفرعية).

- إحصاء كلمات النص:

تقوم برامج معالجة الكلمات بإحصاء عدد الكلمات في النص وتوردها للكاتب عند الطلب، وهي معلومة نحتاج إليها في كتابة البحوث والرسائل العلمية، والواجبات الكتابية المدرسية أحيانًا.

نظرا لهذه الميزات الكثيرة لمعالج الكلمات فإنه وسيلة ممتازة للتدريب على الكتابة في أية لغة.

٢-٢-٤-٢- **التعبيرالكتابي**:

٢-٢-٤-٢-١- التعبير المقيد:

في هذا النوع من التدريبات لا يأتي الدارس بكلمات من عنده. فكل ما عليه فعله هو استخدام الكلمات المعطاة له في التدريب.

هناك تدريبات كتابية عديدة يمكننا إجراؤها بمساعدة برنامج معالجة الكلمات. ومن هذه التدريبات،

- تكملة النصوص الناقصة:

قد تكون هذه النصوص في شكل نصوص قرائية أو حوارات حذفت أجزاء منها، ويقوم الدارس باستكمالها.

- اعادة صياغة النص:

يعطى نص للدارس، ويطلب منه تغييره من شكل إلى آخر(حوار إلى نص سردي) أو تغيير أجزاء منه(تغيير صيغة الأفعال من المضارع إلى الماضي مثلا، أو تغيير الأسماء من مذكر إلى مؤنث أو من المفرد إلى المثنى

أو الجمع) مع عمل التعديلات اللازمة مثل صيغ الصفات أوالضمائر وغير ذلك.

- في المستويات المتقدمة، قد يعطى نص يتسم بالإطناب، ويطلب من الدارس اختصاره. أو يعطى رسالة بأسلوب شبه عامية كي يعيد صياغتها باسلوب أفصح.

٢-٢-٤-٢-٢- التعبير الموجه:

في هذا النوع من التدريبات نقدم للدارس أفكارًا يستفيد منها في كتابة فقرة ما. وقد نزود الدارس بالأفكارعن طريق:

أ- صور مركبة تجمع موضوعات متعددة، ليكتب عنها الدارس (يصف ما يراه.)
ب- صورتين متشابهتين في بعض الوجوه، ويطلب من الدارس المقارنة بينهما.
ت- سلسلة من الصور تحكي قصة ما، ويطلب من الدارس أن يكتب القصة في ضوئها.
ث- مجموعة من الأفكار (العناصر) ويؤلف منها الدارس فقرة مكتوبة. وقد تكون العناصر في شكل أسئلة

يجيب عنها الدارس، ويؤلف نصًا مترابطًا منها.

٢-٢-٤-٢-٣-التعبير الحر:

لا شك أن معالج الكلمات هو أفضل أداة للتدرب على الكتابة الحرة لما لها من ميزات سبقت الإشارة إليها من مسح (حذف) وإضافة وتقديم وتأخير،، إضافة إلى التدقيق الإملائي والمراجعة النحوية.

غير أنه لا بد لنا من إلغاء إمكانيات التدقيق والمراجعة الآلية أحيانا، حتى لا يتأثر إنتاج الدارس بتدخلات الحاسوب(أي يغش الدارس من الحاسوب) من ثم لا يتنبه إليها المعلم ولا يتعرف على قدرات الطالب الحقيقية و أوجه القصور لديه ليعمل على تصحيحها.

ويمكننا تعليق هذه الإمكانيات حتى ينتهي الدارس من كتابته وطباعة ماكتبه. ثم نسمح له باستخدام امكانات التدقيق اللغوي وطباعة النص بعد التنقيح، حتى نتيح للدارس الناضج أن يقارن بين عمله قبل التنقيح وبعده.

٢-٢-٤-٢-٤- كتابة المقالات والبحوث:

إضافة إلى ما سبق، يجدر بنا التنويه إلى الشابكة (الإنترنت) بوصفها مصدرًا غنيًّا ومفيدًا للمعلومات في شتى حقول المعرفة، مما يعين الكاتب والباحث في الحصول على المعلومات التي يحتاج إليها في كتابته أو بحثه.

٢-٢-٤-٢-٥- الألعاب والمسابقات:

يمكن للمعلم أن يقسم الصف إلى مجموعات لكتابة بعض أسئلة المسابقات البسيطة، مثل كتابة ثلاثة جمل عن كل من عدد من الشخصيات الشهيرة ثم يتبادلون الدور لكتابة الإجابات عن هذه الأسئلة.

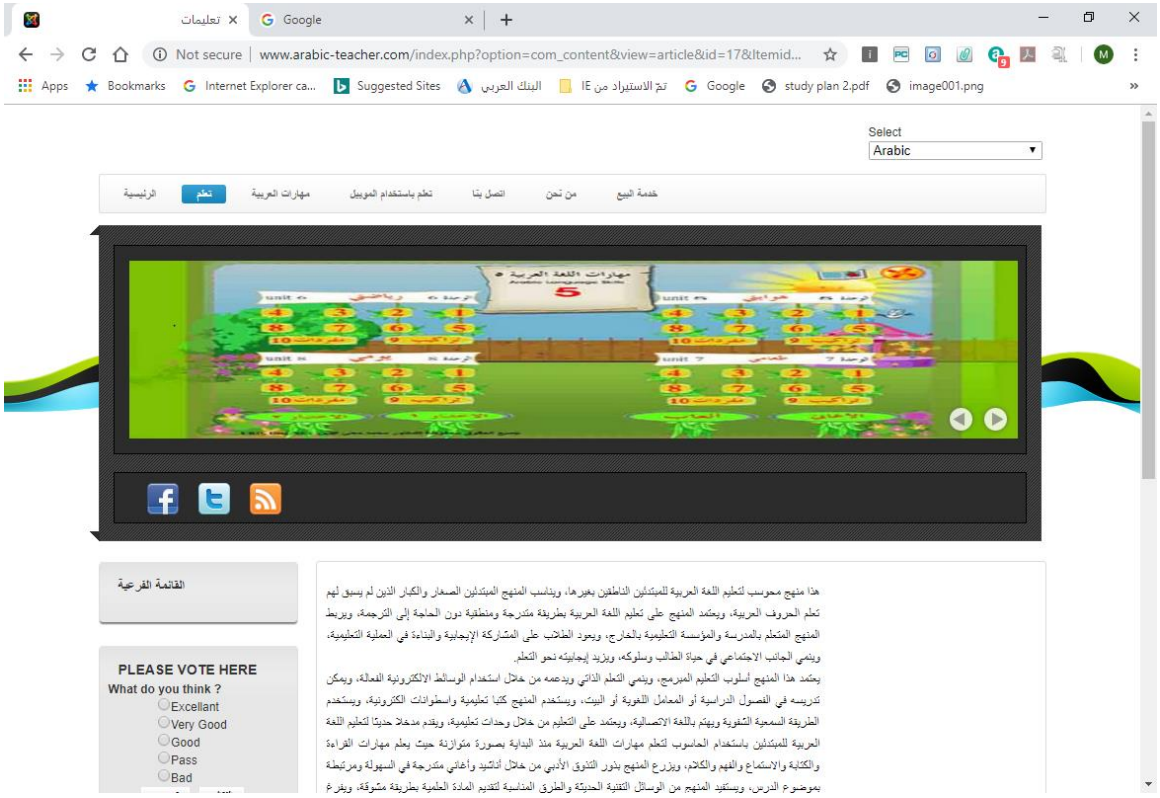
٢-٢-٤-٢-٦- مجلة الصف:

لا شك أن إمكانات الحاسوب المتطورة، خاصة مع برامج(النشر المكتبي) تتيح فرصة ممتازة لتدريب الطلاب على إعداد مجلات للصف أو نادي اللغة أو حتى للمدرسة باللغة الأجنبية التي يدرسونها. وهم أثناء ذلك يتدربون على الكتابة ممزوجة بمتعة التحرير الصحفي والإخراج وغير ذلك من المهارات الجانبية أو المصاحبة للتعبير التحريري.

٢-٣- مثال على حزمة برامج شاملة لتعليم اللغة العربية:

لعل من البرامج الشاملة لتعليم اللغة العربية للناطقين بها وتشتمل على دروس لتعليم الحروف والمهارات والأغاني التعليمية برنامج "معلم العربية" الذي أعدته جامعة بروناي دار السلام بإشراف الدكتور محمد محي الدين في الموقع: www.Arabic-teacher.com

ونورد إحدى شاشات البرنامج المذكور:



٢-٤- التأليف وتصميم الاختبارات:

يمكننا استخدام الحاسوب كبنك للمعلومات للمؤلفين. مثلا إذا كان لدينا عشرة معلمين، وكل واحد منهم يصمم تدريبات في مهارة ما (الاستماع أو الكلام أو القراءة أو الكتابة) أو عنصر لغوي ما (الأصوات مثلًا أو المفردات أو التراكيب)، مع بيان مستوى التدريب (مبتدئ، متوسط، متقدم مثلًا). و نجمع هذه التدريبات في ذاكرة الحاسوب في صورة قاعدة للمعلومات أو البيانات data base فنطلب منه مثلا: "نريد تدريبات

للقواعد أو التراكيب" فيعطيك جميع أنواع التدريبات النحوية المخزونة في ذاكراته. أوتقول "نريد تدريبات للمفردات " فيعطيك جميع أنواع تدريبات المفردات التى وضعت في ذاكراته، وهكذا... ومن المهم هنا أن يذكر المؤلف مستوى كل تدريب (هل هو للمبتدئين أم طلاب المرحلة المتوسطة أم المرحلة المتقدمة). ونطلب من الحاسوب تدريبات للاملاء فيعطيك الحاسوب جميع ما خزن في ذاكرته من التدريبات الموجودة. (انظر المثال في قسم الاختبارات أدناه.)

من البرمجيات المفيدة في التأليف تلك البرمجيات التي تقوم بإحصاء الكلمات في النصوص المختلفة مثل برنامج المكشاف السياقي concordancer، ومن بين تلك البرامج التي تتعامل مع النصوص العربية: aconcorde و Sketech Engine، Ghawwas. فمثل هذه البرمجيات تورد لنا جيع كلمات النص أو مجموعة النصوص مرتبية حسب تكرارها (الأكثر تكرارًا ثم الأقل فالأقل) أو مرتبة ألفبائيا (مع بيان تكرار كل كلمة)، ليكون ذلك مرشدًا للمؤلف، بل ولمصمم الاختبار في اختيار المناسب منها. فيورد المؤلف مثلًا الكلمات الأكثر شيوعًا في المراحل المبتدئة والأقل شيوعًا في المراحل المتقدمة. (لمعلومات حول عملية التأليف ينظر أدناه.)

٢ـ٥ـ استخدامات الحاسوب في الاختبارات:

هناك أربعة استخدامات رئيسة:

أولا: بنوك (قواعد بيانات) الأسئلة
ثانيًا: إعداد أو وضع الاختبارات.
ثالثًا: تطبيق أو تقديم الاختبار مثل ibt (الاختبار المبني على الشابكة)، وكذلك الاختبارات التدريبية على الحاسوب الشخصي.
رابعًا: تقويم أداء المفحوصين أي تصحيح الاختبار
خامسًا: تقويم الاختبار نفسه وتحليله.
ونورد أدناه مزيدًا من التفاصيل حول هذه الاستخدامات.

٢ـ٥ـ١ـ قاعدة بيانات للأسئلة:

تمثل قاعة بيانات الأسئلة مجموعة من الأسئلة مصنفة حسب الموضوع والمستوى مثلا.
مثال على قاعدة بيانات الأسئلة:

من المعروف أن قاعدة البيانات تتكون من (ملفات files) وبكل ملف (سجلات records) يتكون كل منها من (حقول fields).

عنوان الملف: "اختبار عناصر اللغة ومهاراتها". ويشتمل كل سجل على ما يلي؟

حقل ١: السؤال

حقل٢: نوع السؤال (موضوعي، اختيار من متعدد، صحيح وخطأ...)

حقل٣: الموضوع (نحو، صرف، إملاء، فهم مقروء، فهم مسموع...)

حقل٤: الموضوع الفرعي (المطابقة بين الصفة والموصوف، صيغ الأفعال، حروف الجر...)

حقل٥: مستوى السؤال (مبتدىء، متوسط، متقدم...، مستوى أول، مستوى ثان، مستوى ثالث...)

ويتم استرجاع المعلومات (الأسئلة)، وفق رغبات المستعلم (المعلم مثلًا).

عند البحث أو الاستعلام، نطلب مثلا:

الحقل ٣: فهم مقروء، للحصول على جميع الأسئلة الخاصة بفهم المقروء.

الحقل ٣: فهم مقروء+ الحقل ٥: مستوى متقدم، للحصول على جميع أسئلة فهم المقروء للمستوى المتقدم.

الحقل ٤: مطابقة + الحقل ٥: مستوى متوسط، للحصول على جميع أسئلة النحو المتعلقة بالمطابقة وللمستوى المتوسط.

وهكذا دواليك.

٢-٥-٢- إعداد الاختبارات المحوسبة:

هناك عدد من البرمجيات التي تعين واضعي الاختبارات في إعداد الاختبارات المختلفة مثل برنامج Respondus التابع لحزمة WebCT التي تستخدمه جامعة آيوا الأمريكية في إعداد البرامج التعليمية والاختبارات.. وتساعد برمجية Respondus في إعداد الصور المختلفة من الاختبارات (اختيار من متعدد، صحيح وخطأ، ومزاوجة...) (انظر شابل ودوقلاس، ١٤٢٢هـ: صص ٩٧ وما بعدها.)

ومن البرمجيات المتاحة مجانًا على الشابكة برمجية تاوللاختبارات الذي صمم لمساعدة واضعي الاختبارات اللغوية، كما تعين كذلك في تقديم الاختبارات حاسوبيًا. انظر:

http://www.taotesting.com

٢-٥-٣- تقديم الاختبار مع التقويم:

أ- تقديم عام، أي يقدم جميع أسئلة الاختبار لجميع المفحوصين، بغض النظر عن أدائهم.

ب- تقديم الاختبار وفق أداء المفحوصين. وهومايسمى بالاختبار المتكيف حاسوبيا computer adaptive test. ويتم تصميم البرنامج ليتكيف مع استجابات المفحوصين، كأن يقفز بعض الأسئلة إذا كان أداء المفحوص فوق المتوسط.

٢-٥-٤- تقويم الإجابات:

٢-٥-٤-١- تصحيح الاختبارالموضوعي والذي يتم تقديمه ورقيًا أو آليًا. وهذا من أقدم أساليب الاستفادة من الحاسوب، أي تصحيح الاختبارات الموضوعية (خاصة اختبارات الاختيار من متعدد)

من أمثلة البرنامج الذي يقدم الاختبار ويقوم بالتقويم ibt "الاختبار المبني على الشابكة internet" للتوفل. كما أن هناك برامج للتدريب على الاختبارات (التوفل مثلا) في أقراص مضغوطة، تقدم الاختبار وتقوم بعملية التقويم. انظر أيضًا: https://www.transparent.com/language-resources/tests.html وهي مؤسسة تقدم اختبارات للتقويم الذاتي على الشابكة في عدد من اللغات.

٢-٥-٤-٢- تقويم اختبار الإنتاج اللغوي (الكلام والكتابة)

من أمثلة البرامج المتطورة لتقويم الإجابات المكتوبة، برنامج يسمى e-rater "المقوم الإلكتروني " لتقويم المقالة (أي التعبير التحريري) بطريقة آلية، أي دون تدخل بشري. وهناك برنامج آخر اسمه "كرايتيريون Criterion" لتقويم المقالات، أي كتابات الطلاب بطريقة آلية.

أما الإنتاج الشفوي فقد أنشأت له أوردنيت، إنك Ordinate, Inc) برمجية تعرف باسم SET-١٠ الذي يتواصل به المفحوصون هاتفيًا، لتتم عملية تقويم أدائهم الشفوي. (انظر شابل ودوقلاس، ١٤٢٢هـ: صص ٥٤-٥٨.

(انظر: كارول شابل ودان دوقلاس (٢٠٠١) وترجمة سعد بن على القحطاني. **تقييم اللغة باستخدام تقنية الحاسوب.** الرياض: دار جامعة الملك سعود للطباعة والنشر.)

ولتقويم الأداء الشفوي كذلك يستخدم برنامج اللغة الأمريكية في جامعة كولومبيا في نيويورك (برمجية آول للاختبارا ت) الذي يستخدم في إعداد الاختبارات اللغوية وتقويمها. انظر:

testing software-https://www.owlts.com/owl-testing-software.aspOWL

انظر أيضًا: http://www.icls.edu/language-skills-assessment/ وانظر "التقييم الآلي" لفهمي (٢٠١٩)

٢-٥-٤-٥- تقويم الاختبارات:

هناك حزم حاسوبية تقوم بعملية تحليل البنود (الأسئلة) من حيث الصدق المقارن والثبات والسهولة والصعوبة ومن حيث التمييز.

٢-٦- تقويم المواد التعليمية:

منذ مدة طويلة بدأ الدكتور ماكي Mackeyفي استخدام الحاسوب في تحليل المواد التعليمية: لدراسة مثلا نسبة ورود نقاط معينة، أو ورود تراكيب معينة في الكتاب. فالحاسوب يفيد هنا في التحليل، ومن ثم يفيد في التأليف،حيث يعين المؤلف في تحليل المادة التي يعدها، فيتبين له إذا كانت هناك مثلا كلمات جديدة أو تراكيب لم ترد أو أنه لم يكرر ظاهرة معينة بدرجة كافية، إلى غير ذلك مما يجب مراعاته عند إعداد المادة التعليمية.

يمكننا لهذا الغرض أن نستعمل برمجيات مختلفة، مثل معالج الكلمات ويرد word أو برنامج قواعد البيانات مثل أكسس Access وبرامج المكشاف السياقي concordancers التي ذكرناها سابقًا. وهذا يحتاج إلى ورقة خاصة لمناقشة تفاصيل ليس هنا مكانها.

٢-٧- الشابكة (الإنترنت) في تعليم اللغة:

إضافة إلى ما أشرنا إليه أعلاه من استخدامات الشابكة في التدريب على الاستماع والقراءة والقواعد والمفردات والكتابة، هناك استخدامات أخرى يجدر بنا الإشارة إليها. (انظر الدراسة التالية لمناقشة ممتازة للموضوع.)–!٢٠١١!June !Online!ToolsforLanguage!Teaching !Volume!١٥,!Number!١ Jeong@Bae!Son!

ويصنف الباحث سون Son أدوات الشابكة إلى اثنتي عشرة فئة: منظومات إدارة التعلم أوالمادة التعليمية، التواصل، العوالم الحية والافتراضية، الشبكات الاجتماعية، المدونات، الويكي، العرض، التشارك في المصادر، إعداد مواقع الشبكات، إعداد التدريبات على الشابكة، أدوات البحث على الشابكة، المعاجم،

والمكشافات السياقية. ولقد سبق أن تطرقنا إلى بعض من هذه الأدوات في أجزاء مختلفة من دراستنا هذه.

ونوجز فيما يلي بعض هذه الاستعمالات:

٢-٧-١- الشابكة أداة لدراسة اللغة:

من المعروف أن هناك كثيرًا من البرامج الحاسوبية على الشابكة لتعليم اللغات المختلفة. هناك قائمة لكثير من المواقع على الشابكة (الإنترنت) لتعليم اللغة العربية بوصفها لغة أجنبية، أعده محمود إسماعيل صالح ضمن دليل للسلاسل والمواقع المعروفة. (انظر صالح (٢٠٢٠) في قائمة المراجع.)

٢-٧-٢- الشابكة أداة للتواصل:

من الأساليب المفيدة في تعليم وتعلم اللغات التواصل مع الناطقين الأصليين للغة. والشابكة وسيلة ممتازة لذلك، حيث يمكن للدارس التراسل مع بعض الناطقين باللغة سواء في صورة تراسل اجتماعي، أو عن طريق تبادل التعليم إذا كان الطرف الآخر ممن يتعلمون لغة الدارس، فيقوم كل منهما بمساعدة الطرف الآخر في تعلم لغته. كذلك تتيح لنا الشابكة مشاركة صفوف تعليم اللغة في مناطق مختلفة من العالم في بث مباشر والتحاور مع المعلمين والدارسين في تلك الصفوف.

٢-٧-٣-الشابكة أداة للتعاون:

من الأساليب الحديثة في تعليم اللغات ما يسمى بالتعلم التعاونيcooperative learning ولا شك أن الشابكة خير وسيلة لهذا النوع من التعلم، حيث يمكن لأي عدد من الدارسين التعاون مع بعضهم البعض في العملية التعليمية، سواء أكان في صورة المشاركة في أداء بعض المهام (مثل كتابة فقرة أو مقال) أم في صورة التعاون في مراجعة أو تقويم أعمال بعضهم البعض،

وذلك بإشراف المعلم أو بدونه.

(لمزيد من المعلومات حول هذا الموضوع انظر كتاب *The Internet in the Language Classroom* ودراسة سون Son المشار إليها أعلاه).

٢-٧-٣- الشابكة بوصفها مصدرًا للمعلومات:

لاشك أن كل من استخدم الشابكة، خاصة google، يدرك الفائدة غير المحدودة من الشابكة في الحصول على شتى أنواع المعلومات اللغوية وغير اللغوية من مصادر ومواقع مختلفة. بل ربما اقتصر استعمال الكثير منا للشابكة على هذا الغرض.

٢-٨- بعض مشكلات تعليم اللغة باستخدام الحاسوب وحلول مقترحة:

سنتحدث أدناه عن بعض من أهم "المشكلات" الكامنة في استخدام الحاسوب في تعليم اللغات. وسنذكر أدناه المشكلة (م) ثم نعطي بعض الحلول المقترحة أو المحتملة (ح).

م- يتطلب استخدام الحاسوب إجادة استخدام لوحة الطباعة (أي مثل الضرب على الآلة الكاتبة)، حيث إن المبتدئ يضيع وقتا كبيرا في البحث عن المفاتيح المناسبة لما يريد كتابته. وكثير من البرامج تحاسب الدارس على الوقت الذي يقضيه في الإجابة عن سؤال معين مثلا -فنجد أن الدارس قد يعرف الإجابة في الحال. ولكنه يحتاج الى وقت طويل لكي يدخلها في الحاسوب.

ح-١- هناك حلول جزئية لهذه المشكلة في الوقت الحاضر، فمنها استخدام الفأرة وعمليات القص والنسخ واللصق (cut, copy, past) المتاحة في جميع برمجيات تنسيق النصوص. ومنها اختصار إجابة الدارس إلى أقصى حد ممكن ، كأن يكتب حرفا واحدا مثلا بدلا من جملة كاملة، وذلك كما في أسئلة الاختيار من متعدد.

مثال:

أكمل الجملة التالية بكتابة رقم أو حرف الاجابة الصحيحة: -

يذهب التلاميذ في الصباح الى

أ- مدارسهم

ب- مصانعهم

ج-.مكاتبهم

كل ما على الدارس أن يفعله هو كتابة الحرف (أ) للإجابة عن مثل هذا السؤال. كما أنه يمكن للدارس أن ينتقل بالمشيرة curserإلى موقع الإجابة الصحيحة ثم يضغط على مفتاح الإدخال enter.

ح - ٢ - قبول استجابة الدارس شفويا. غير أن هذا لايزال في مرحلة التجريب والتطوير، فالحاسوب لم يصل بعد إلى الدرجة التي تمكنه أن يقبل بها إجابات حرة بأصوات مختلفة في جميع اللغات.

م- تقديم السياقات الطبيعية من صور ومواقف حية لاستخدام اللغة.

ح- ربما يمكننا حل هذا جزئيا باستخدام الاسطوانات المرئية(videodisc) التي تستخدم أشعة الليزر في تشغيلها، ونظام اليوتيوب YouTube المنتشرة على الشابكة.

(تحويل المحتوى الى صور ثابتة أو متحركة) يشاهدها الدارس على شاشة الحاسوب مع الصوت أو بدونه.

-م- التهجئة الخاطئة والإجابة الصحيحة: قد يرفض الحاسوب بعض الإجابات لمجرد خطأ طفيف في تهجئة بعض الكلمات.

ح- توسيع مدى الاستجابات المقبولة، أو التساهل في عدد من الأخطاء، كأن تعطى تعليمات للحاسوب بقبول الكلمة أو الجملة التي تختلف عما في البرنامج في حرفين أو ثلاثة مثلا وحديثًا نجد حلولًا تستفيد من الذكاء الاصطناعي.(AI) Artificial Intelligence انظر"التقييم الآلي" لفهمي (٢٠١٩) وعبدالغني (٢٠١٩) والحايك (٢٩١٩).

م- رفض الإجابة الصحيحة المخالفة لما في البرنامج. قد يأتي الدارس بإجابة صحيحة ويعتبرها الحاسوب خطأ لأنها غير مخزنة في ذاكرته.

ح- كتابة عدد كبير من الإجابات المحتملة. فكلما زادت هذه زاد احتمال قبول الحاسوب لإجابة الدارس.

مثلا:

العبارة (كيف الحال؟) قد يرد عليها بتعبيرات كثيرة مثل: -

(أنا بخير) (أنا بخير الحمد لله) (الحمد لله) (لابأس)(الحمد لله، وأنت كيف حالك؟) وهكذا. وبذلك يزداد احتمال قبول اجابة الدارس.

م- صعوبة تقديم الجانب الشفوي للغة:

ح-١- الحل الأمثل هو الذي بدأ ينتشر حاليا، ألا وهو الأقراص المضغوطةCopact Disc (CD Rom) وكذلك الأقراص البصرية video disc لتسجيل الصوت عليها (مع الصور أو بدونها). وتمتاز هذه بإمكانية التنقل عليها بسرعة عن طريق أشعة الليزر. ومؤخرًا انتشر نظام MP٣ لرقمنة الكلام، لتسجيله صوتيًا. هذا إضافة إلى اليوتيوب You-Tube الذي أشرنا إليه أعلاه.

ح-٢- طورت بعض الشركات نظاما لتقديم الكلام الاصطناعي synthetic speech ما يعرف عادة. فهناك برمجيات المركب للأصوات اللغوية speech synthesizer، أو تحويل النص إلى كلام speech-to-text فتنطق بالكلمات أو الجمل المكتوبة، ولكن بشكل غير طبيعي، من حيث غياب التنغيم الطبيعي للكلام (وهو أمر هام جدا عند تعليم اللغة الاجنبية).

ح-٣- نلاحظ مؤخرا وجود قواميس إلكترونية يتم فيها تسجيل النطق الطبيعي. فعند اختيار كلمة عربية مثلا وطلب المقابل الانجليزي لها يتيح لنا هذا النوع من القواميس قراءة الكلمة الأجنبية وسماع نطقها كذلك، والعكس صحيح أيضا، أي نختار أو ندخل كلمة أجنبية ونطلب مقابلتها بالعربية حيث تظهر لنا الكلمة مصحوبة بنطقها. وقد نختار كلمة عربية ما فنسمع نطقها ونشاهد كتابتها.

م- تهيُّب كثير من الناس من الأجهزة غير المألوفة، ولو أن هذا لم يعد مشكلة خاصة مع الأجيال الجديدة من المعلمين والدارسين.

ح -الحل الوحيد هنا هو من خلال:

أ- شرح أعمال الحاسوب وإجراءات الدرس بالتطبيق العملي.

ب- التقليل من الأعمال التى يقوم بها الدارس أثناء استخدامه للحاسوب (كما في المثال المذكور في ٢ آنفا).

ج- إعطاء التعليمات الواضحة والبسيطة للدارس في ثنايا البرنامج، وذلك إما صوتيا أو كتابة على الشاشة أمامه، خطوة خطوة، وذلك من خلال قوائم الخيارات الجاهزة menusالتي ماعلى الدارس سوى الاختيار منها بمجرد وضع المشيرة عليها.

م- ضرورة استخدام لغة وسيطة (أي لغة غير اللغة التي يدرسها الطالب) في أى برنامج للتعليم الفردى للغة الاجنبية.

خ- علاج هذه المشكلة يتم عن طريق التقليل من دور اللغة الوسيطة تدريجيا مع تقدم الطالب في اللغة الاجنبية - حتى نتخلص منه نهائيا في المراحل المتقدمة. عليه فإن البرامج المعدة للدارسين المتقدمين لا تتطلب استخدام لغة وسيطة. ولا تستعمل لغة أخرى غير اللغة المدروسة (الهدف) إلا فى حالة الترجمة إن كانت مقصودة لذاتها.

٢-٨- مشكلة التأليف وإعداد الدروس المحوسبة:

المشكلة الجوهرية في استخدام الحاسوب أحيانًا التأليف. غير أن هذا ينطبق على التأليف المتطور الذي يتيح الاستفادة من الإمكانات المختلفة للحاسوب مثل التغذية الراجعة (أي إعلام الدارس مثلًا بصحة إجابته من خطئه فيها)، والإحالة آليًا إلى الملفات أو المراجع المختلفة باستخدام النص الخفي hypertext، والربط بالصوت والصورة (ما يسمى الوسائط المتعددة multimedia) ومتابعة سير الدارس وتقويم أدائه العام...إلى غير ذلك.

من ناحية أخرى، نجد أن المعلم يستطيع أن يعد كثيرًا من الدروس والتدريبات التي ذكرناها أعلاه دون الحاجة إلى أية معلومات في البرمجة. بل يمكنه استخدام برنامج معالج الكلمات مثل word فقط. أما إذا أرا د أن يعد مواد متطورة فلا بد له من الاستعانة بمبرمج حاسوبي لتصميم المادة التعليمية.

من الحلول المتاحة البرمجيات الخاصة بالتأليف، مثل منظومات التأليف authoring systems. وهناك عشرات البرمجيات ا (المتاح بعض منها مجانا على الشابكة. وكل ما على المؤلف أن يفعله أن يكتب

عبارات كالتالي لتظهر له قوائم بعشرات البرمجيات (كلمات البحث: authoring system, language teacher أو ما شابه ذلك)، والبرنامج معد مسبقا لمساعدة المؤلفين. فتعليمه واحدة يكتبها المؤلف تنتج منه مئات التعليمات التى تسجل في الحاسوب. أي أن منظومة أو برنامج التأليف يترجم التعليمات التي يكتبها المؤلف للحاسوب بلغته كي ينفذها. إضافة إلى ذلك يقدم البرنامج خيارات عديدة للمؤلف لتصميم دروسه على الحاسوب.

فالمشكلة الأساسية تكمن هنا في خيال المؤلف وفي خيال وإمكانات المبرمج (أحيانا). فإذا اجتمع هذان ربما أمكننا تقديم ٨٠% من دروس اللغة أو أكثر باستخدام الحاسوب. فالآن أصبحت لدينا الصورة والصوت وإمكانات التحكم السريع وإمكانات الاستجابة السريعة والتحليل السريع. وهي جميعا متاحة في الحاسوب الذي أصبح نفسه متاحا لكثير من الناس في صورة الحويسبات الشخصية والمحمولة laptops واللويحات tablets والهواتف الذكية smart phones وغيرها ذوات القدرات العالية والأثمان المنخفضة نسبيا.

المراجع العربية

بوعزاوي، المصطفى وآخرون (٢٠١٦) **التقنية في تعليم العربية للناطقين بغيرها**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

جمعة، عمرو (٢٠١٦) **تقنيات اللغة العربية الحاسوبية: معايير التقييم ورؤى التطوير**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

الحاج، يحي محمد (٢٠١٥) "تقنيات التعرف الآلي على الكلام المنطوق وتطبيقاتها في القرآن الكريم: واقع وطموح" في: العريان، يوسف سالم (٢٠١٥): ٢٤٥-٢٨١.

الحايك، أحمد (٢٠١٩) "التعليم العميق وتطبيقاته المرتبطة باللغة العربية" في: السعيد (تحرير، ٢٠١٩) ١٤٣-١٦٣.

حسن، محمد أحمد عبده (٢٠١٣) الاحتياجات التكنولوجية لمعلمي اللغة العربية للناطقين بغيرها. مركزالشيخ زايد لتعليم العربية لغير الناطقين بها.

تاريخ الوصول ٢٠١٩/٨/١٦.

رشوان، محسن و السعيد، المعتز بالله (٢٠١٩) **تطبيقات أساسية في المعالجة الآلية للغة العربية**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

السعيد، المعتز بالله (٢٠١٩) **العربية والذكاء الاصطناعي**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

شابل، كارول، ترجمة سعد القحطاني (٢٠١١) **تطبيقيات الحاسب الآلي في اكتساب اللغة الثانية**. الرياض: جامعة الملك سعود.

شابل، كارول ودان دوقلاس (٢٠٠١) وترجمة سعد بن على القحطاني. **تقييم اللغة باستخدام تقنية الحاسوب.** الرياض: دار جامعة الملك سعود للطباعة والنشر

شاتوك، كاي (تحرير، ٢٠١٧) **ضمان الجودة في التعليم عبر الإنترنت: الممارسات والعمليات على مستوى التدريس والموارد والبرامج**. ترجمة محمد إبراهيم الزكري. الرياض: معهد الملك عبدالله للترجمة والتعريب.

شتات، خالدة عبد الرحمن (٢٠١٠) "تعليم اللغة العربية بوساطة الحاسوب في الصفوف الأربعة الأولى: الواقع والمأمول". في **مجلة مجمع اللغة العربية الأردني.** *٢٧ تشرين الأول ٢٠١٠م*

https://www.majma.org.jo/res/seasons/٢٨/٢٨-١٧.do

صالح، محمود إسماعيل (١٤٣٦هـ) "المدونات اللغوية وكيفية الإفادة منها" في العصيمي، صالح (١٤٣٦هـ): ١٧-٩٣.

صالح، محمود إسماعيل (٢٠٢٠) **دليل السلاسل التعليمية والمواقع لتعليم العربية**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

صيني، د. محمود اسماعيل، ود. على محمد القاسمى (١٩٨٠م) السجل **العلمى للندوة العالمية الاولى لتعليم العربية لغير الناطقين بها الجزء الثاني، المعلم، الكتاب، الطريقة، الوسائل.** الرياض: عمادة شئون المكتبات بجامعة الرياض.

تاريخ الوصول: ٢٠١٩/٨/١٦

عبدالغني، نعيم محمد (٢٠١٩). الذكاء الاصطناعي وتعليم اللغة العربية: نحو منصّة تعليمية متكاملة. في: السعيد (٢٠١٩): ١٧٦-٢٠٩.

عبد المجيد، عواطف حسن علي: تعليم اللغة العربية بوساطة الحاسب الآلي.

www.alarabiahconference.org/.../conference_research-٦

عبده، شريف (٢٠١٩) "تطبيقات معالجة اللغة العربية في مجال التعليم" في رشوان، محسن و السعيد، المعتز بالله (٢٠١٩)، ٩٣-١١٩.

عبود، فكتورين "استخدام العقل الالكترونى كأداة في تعليم العربية كلغة أجنبية " في صينى والقاسمى (١٩٨٠): ٣١٥-٣٢٧.

العتيبي، هند بنت مطلق (٢٠١٧) "المعالجة الآلية للمدونات المتوازية واستخداماتها في تعليم اللغات وتدريب المترجمين" في: المجيول، سلطان (٢٠١٧): ١٦٩-١٩٧.

العريان، يوسف سالم (تحرير ٢٠١٥) **الحرف العربي والتقنية: أبحاث في حوسبة العربية**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

العصيمي، صالح بن فهد (تحرير، ١٤٣٦هـ/ ٢٠١٥م) **المدونات اللغوية العربية: بناؤها وطرق الإفادة منها**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

فرح، د. قيصر."العقل الالكترونى في تعليم اللغة العربية لغير العرب" في صيني والقاسمى (١٩٨٠م) ٣٢٩-٣٣٧.

فهمي، علي (٢٠١٩) "التقييم الآلي" في: رشوان والسعيد (٢٠١٩)، ١١٩-١٤٥.

الفيفي، عبدالله بن يحي (٢٠١٧) "المعالجة الألية لمدونات المتعلمين: تحليل الأخطاء في المدونة اللغوية لمتعلمي اللغة العربية" في: المجيول، سلطان بن ناصر(٢٠١٧): ١٣٥-١٦٨.

القاسمى، على (١٣٩٩هـ) " استخدام العقل الالكترونى في التربية" في علي القاسمي: **اتجاهات حديثة في تعليم العربية للناطقين باللغات الأخرى.** الرياض: عمادة شئون المكتبات بجامعة الرياض

المجيول، سلطان بن ناصر (٢٠١٧) **لغويات المدونة الحاسوبية: تطبيقات تحليلية على العربية الطبيعية**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

ولهاصي، نجيدة: إمكانات استخدام الحاسوب في تعلَم وتعليم اللغات http://www.inst.at/trans/٢٢/

تاريخ الوصول: ٢٠١٩/٨/١٦.

مراجع أجنبية:

Barth, Melisssa E (١٩٨٨) *Strategies for Writing with the Computer*. New York: McGraw-Hill.

Brierley, W and L R Kemble (eds., ١٩٩١) *Computers as a Tool in Language Teaching*. New York: Ellis Horwood.

Bush, M. D. (٢٠٠٨). Computer---assisted language learning: From vision to reality? *CALICO Journal, 25*(٣), ٤٤٣---٤٧٠.

Butler, C S (ed., ١٩٩٢*) Computers and Written Tex*ts. Oxford: Basil Blackwell.

Chapelle, C., & Jamieson, J. (٢٠٠٨). *Tips for teaching with CALL: Practical approaches to computer--assisted language learning*. White Plains, NY: Pearson Education.

Chun, D. M. (٢٠٠٧). Come ride the wave: But where is it taking us? *CALICO Journal, 24*(٢), ٢٣٩---٢٥٢.

Davis, G and J Higgins (١٩٨٥) *Using Computers in Language Learning: A Teacher's Guide.* London: Multiplex Techniques Ltd.

De Szenkeffy, John (٢٠٠٥) *A Practical Guide to Using Computers in Language Teaching***.** Ann Arbor, Michigan: University of Michigan Press.

Dunkel, Patricia (ed., ١٩٩١) *Computer Assisted Language Learning and Testing:Research Issues and Practice. Rowley, MA: Newbury Houses*

Garrett, N. (٢٠٠٩). Computer---assisted language learning trends and issues revisited: Integrating innovation. *The Modern Language Journal, 93* (Focus Issue), ٧١٩---٧٤٠.

Godwin---Jones, R. (٢٠٠٩). Personal learning environments. *Language Learning & Technology, 13*(٢), ٣---٩. Retrieved May ١٨, ٢٠١١, from http://llt.msu.edu/vol١٣num٢/emerging.pdf.

Godwin---Jones, R. (٢٠١٠). Literacies and technologies revisited. *Language Learning & Technology, 14*(٣), ٢---٩. Retrieved May ٢٠, ٢٠١١, from http://llt.msu.edu/issues/october٢٠١٠/emerging.pdf.

Hardisty, D. and S Windeat (١٩٨٩) *CALL (Resource Books for Teachers Series)* Oxford: OUP.

Higgigns, J and T Johs (١٩٨٤) *Computers in Language Learning.* London: Collins ELT.

Howatt, A P R (١٩٦٩) *Programmed Learning and the Language Teacher.(Longman Handbook for Language Teachers).* London: Longman.

Jones, C. and S Fortescue (١٩٨٧) *Using Computers in the Language Classroom*. London: Longman.

Kenning, M J and M-M Kenng (١٩٨٣*) An Introduction to Computer Assisted Language Teaching*. Oxford: OUP.

Lareav, P and E Vockell (١٩٨٩*) The Computer in the Foreign Language Curriculum.* New York: Mitchell Publishing, Inc.

Leech, G and C.N Candlin (eds., ١٩٨٦*) Computers in English Language*

Levy, M. (٢٠٠٩). Technologies in use for second language learning. *The Modern Language Journal, 93* (Focus Issue), ٧٦٩---٧٨٢.

Mathews, Dorothy (ed. ١٩٨٥) *The Computer and English Teaching.* Urbana: Illinois Assn of Teachers of English.

Meskill, C., & Anthony, N. (٢٠١٠). *Teaching languages online*. Bristol, UK: Multilingual Matters.

Pennington, Martha C V. Stevens (eds., ١٩٩٢*) Computers in Applied Linguistics: An International Perspective.* Avon, England: Multimedia Matters.

Schwartz, E and E Vockell (١٩٨٨*) The Computer in the English Curriculum.* New York: Mitchell Publishing.

Son, J.---B. (٢٠٠٢). Online discussion in a CALL course for distance language teachers. *CALICO Journal, 20*, ١٢٧---١٤٤.

Son, J.---B. (٢٠٠٤). *Internet--based language instruction: Study book*. Toowoomba, Australia: Distance Education Centre, University of Southern Queensland.

Son, J.---B. (٢٠٠٥). Exploring and evaluating language learning Web sites. In J.---B. Son & S. O'Neill (Eds.), *Enhancing learning and teaching: Pedagogy, technology and language* (pp. ٢١٥---٢٢٧). Flaxton, Australia: Post Pressed.

Son, J.---B. (٢٠٠٧). Learner experiences in Web--based language learning. *Computer Assisted Language Learning, 20*(١), ٢١---٣٦. TESL-EJ ١٥,١, June ٢٠١١ On the Internet/Son page ١٢

Son, J.---B. (٢٠٠٨). Using Web---Based Language Learning Activities in the ESL Classroom. *International Journal of Pedagogies and Learning, 4*(٤), ٣٤---٤٣.

Son, J.---B. (٢٠٠٩). Using Web---based portfolios in CALL teacher education. In J.---B. Son (Ed.), *Internet--based language learning: Pedagogies and technologies* (pp. ١٠٧---١١٨). APACALL Book Series Volume ٢. Raleigh, NC: Lulu.

Son, Jeong-Bae. (٢٠١١) Online Tools for Language Teaching, *Electronic Journal for English as a Second Language,* June ٢٠١١, Volume١٥, Number ١.

Stanley, G. (٢٠١٣) Language *Learning with Technology: Ideas for integrating technology in the classroom*. Cambridge: CUP.

Vockell, E. L., and Schwartz, E. (١٩٩٢). *The Computer in the Classroom* (٢nd Ed.). New York: McGraw-Hill.

Warschauer, M. (٢٠١٠). New tools for teaching writing. *Language Learning & Technology, 14*(١), ٣---٨. Retrieved May ١٨, ٢٠١١, from http://llt.msu.edu/vol١٤num١/commentary.pdf.

Whitaker, B T, E Schwartz and E.L Vockell (١٩٨٩*) The Computer in the Reading Curriculum.* New York: Mitchell Publishing and McGraw-Hill, Inc.

الترجمة والتعريب

التقنية في خدمة الترجمة:
طرق الاستفادة والتقويم

المقدمة:

شهد النصف الثاني من القرن العشرين والعقد الأول من القرن الحادي والعشرين تطورا مذهلا في مجال الحواسيب والبرمجيات الحاسوبية كما وكيفا، وقدمت بذلك إمكانات وخدمات كبيرة للعاملين في شتى المهن وحقول المعرفة. ولا شك أن الترجمة لقيت الكثير من العناية في هذا المجال، بدءا بالمحاولات الأولية في العقد السادس من القرن الماضي لترجمة النصوص العلمية من الروسية إلى الإنجليزية.

وقد شاهدنا في الآونة الأخيرة تفجرا في أنواع المعينات الحاسوبية للمترجمين، على مستويات مختلفة، تبدأ بما يسمى بالترجمة الآلية (أو الترجمة بمعاونة الحاسوب) وانتهاء بمنسقات النصوص والإملاء الآلي automatic dictation ومابينهما من ذاكرات الترجمة وبنوك المصطلحات الآلية terminology data banks والمعاجم المحوسبة (أحادية اللغات وثنائيتها).

وسنتحدث فيما يلي بإيجاز عن كل من هذه الوسائل المعينة للمترجمين.

١- الترجمة الآلية:

يعرف استخدام الحاسوب أداة للترجمة بالترجمة الآلية machine translation، ويميل البعض إلى تسمية يرونها أكثر ملاءمة، وهي الترجمة بمعاونة الحاسوب computer aided translation.

سأحاول أن أقدم مثالا على طريقة عمل إحدى النظم المعروفة للترجمة الآلية ، وهو نظام توم TAUM (الترجمة الآلية لجامعة مونتريال) وهومن إنتاج جامعة مونتريال بكندا، ويستخدم حاليا أحد نواتج هذا النظام - ميتيو Meteo - لترجمة التقاريروالتنبؤات الجوية من الإنجليزية إلى الفرنسية في كندا.

المرحلة الأولى:

نجد في هذه المرحلة أنه يتم إدخال النص المراد ترجمته، وذلك إما عن طريق لوحة المفاتيح في الحاسوب أو ينقل إلى الحاسوب عن طريق الذاكرة المتنقلة (يواس بي USB مثلًا) أو أية وسيلة أخرى، مثل قارئة المحارف البصرية optical character reader. وبعد عدة خطوات يخرج لنا النص في اللغة الهدف أي اللغة المترجم إليها.أما الخطوات التي تتم في الترجمة الالية فهي:

١-يقوم المترجم البشري بعملية مراجعة مبدئية للنص المراد ترجمته ؛ فالحاسوب لايميز مثلا بين الكلمة الحقيقية وأسماء العلم.فلابد للإنسان أن يوضح أن هذا اسم علم مثلا و لايترجم. وأن هذه الأشكال ليست نصوصا للترجمة. وهذه الخطوة تسمى بعملية التحرير المسبق.

٢-ثم يقوم الحاسوب بعد ذلك بتحليل صرفي للنص للتعرف على أجزاء الكلمات ومكوناتها وأقسامها، وهذا يسير وفق جداول تعطى للحاسوب.

٣- بعد أن يقوم الحاسوب بعملية التحليل الصرفي يقوم بعملية البحث في المعجم (مثلا حينما نجد كلمة schools لابد أن تحلل هذه صرفيا فإذا بحثت عن schools في المعجم لن تجدها لأن الكلمة تتكون من school+s بمعنى مدرسة +علامة الجمع).إذاً لابد من التحليل الصرفي قبل النظر في المعجم. و يقوم الحاسوب بالتحليل الصرفي ثم بالنظر في المعجم المزود به. ثم بعد ذلك يطبع لنا الكلمات التي لم يجدها في المعجم ليخبرنا بأن هذه الكلمات ليست موجودة في قاموسه. عند ذلك يقوم المترجمون بتزويد الحاسوب بتلك الكلمات التي لم تكن مخزونة في ذاكرته وبمقابلاتها، وكذلك بكافةالمعلومات الصرفية والنحوية اللازمة.

٤- يقوم الحاسوب بالبحث في القاموس مرة اخرى.

٥-بعد ذلك يقوم الحاسوب بالتحليل النحوي للنص (أي إعراب الجمل لتحديد الوظيفة النحوية لكل كلمة.

المرحلة الثانية:

تبدأ خطوات الترجمة. وهنا تتم خطوتان حيث:

٦- يقوم الحاسوب بنقل الكلمات حسب قاموس الترجمة فينقل من اللغة الأصل إلى اللغة الهدف، مستفيداً في ذلك من المعجم ثنائي اللغة المتاح له.

٧- ثم يقوم بعملية ترجمة التراكيب: يضع المقابلات التركيبية في اللغة المترجم إليها، مثلا الفاعل والمفعول به والمضاف والمضاف إليه والصفة والموصوف وغير ذلك . أي يقوم الحاسوب بعمليات مناقلة نحوية

بين اللغتين.

المرحلةالثالثة:

تأتي هنا مرحلة التركيب أو التوليف، أي التعبير باللغة الهدف. ويتم ذلك في خطوتين كذلك:

٨- التوليف النحوي، أي تطبيق قواعد النحو في اللغة الهدف، مثل وضع الفعل في بداية الجملة العربية، وليس البدءبالفاعل كما في الانجليزية مثلا، ووضع المضاف إليه بعد المضاف والصفة بعد الموصوف (في اللغة العربية)، بدلاً من ترتيبها العكسي في كثير من اللغات الأخرى.

٩- التوليف الصرفي، وذلك بتطبيق القواعد النحوية والصرفية في اللغة الهدف مثل ملاحظة أن كلمة "سامع+جمع+مفعول به " تصبح "سامعين" بينما "كتب+مضارع+فاعل +جمع مذكر" تصبح "يكتبون" وهكذا.

إذاً هناك ثلاث مراحل أساسية: هي التحليل ثم النقل (أي الترجمة الأولية) ثم تكوين الجمل في اللغة المترجم إليها.

هذا باختصار مثال لما يمكن أن يقوم به الحاسوب في الترجمة. ولكن الموضوع ليس بهذه البساطة. فكما ذكرنا سابقا هناك مشكلات كثيرة تعترض الحاسوب في الترجمة. لذلك نجد أن هناك ثلاثة أساليب مختلفة ظهرت في الترجمة الآلية هي:

أ-الترجمة الآلية مع تحرير لاحق، أي مراجعة بشرية بعد الترجمة الآلية.

ب-الترجمة مع التحرير السابق، بمعنى أن الإنسان يحرر النص المراد ترجمته. مثلا نبسط الجمل المعقدة. والكلمات التي لها معاني كثيرة نحدد معناها المطلوب وهكذا ، أي أننا نعدل النص بحيث يستطيع أن "يفهمه" الحاسوب، وتسمى هذه اللغة المقبولة للآلة (MAL (Machine Acceptable Language.

جـ- ثم هناك نوع ثالث يسمى بالترجمة التحاورية interactive وهي مثال للتعاون بين الحاسوب وبين المترجم البشري، كما سنوضح أدنا.

أ- الترجمة الآلية مع التحرير اللاحق post-editing:

ثبت من التجربة أن الترجمة الآلية حتى للنصوص العلمية والتقنية تتطلب مراجعة بنسبة لاتقل عن ٢. إلى ٤. % أحيانا ، حتى تصبح الترجمة مقبولة للنشر. وهذا التنقيح يشمل تعديلات من حيث المعجم أو الكلمات و من حيث التراكيب وغير ذلك. أما إذا كانت الترجمة لغرض الاطلاع العام والتعرف على محتوي النص

بصورة إجمالية فقد لانحتاج إلى تنقيحها. و وهذا هو أسلوب الترجمة التي يتبعها نظام سيستران Systranالذي يستخدم في أمانة المجموعة الاقتصادية الأوربية EEC في لوكسمبورق.كما أن هذا هو الأسلوب الذي تتبعه القوات الجوية الأمريكية مستخدمة نظام سيستران كذلك.

جدير بالذكر أن بعض الهيئات كانت تعطي الترجمة الأولية للمتخصصين (مثلا علماء الفيزياء النووية) للإطلاع عليه، ثم تعيد ترجمة النص بشريا إن وجد أنه يستحق ذلك.

ب- الترجمة مع التحرير السابق pre-editing:

أما الاتجاه الثاني، وهومايسمى بالترجمة الآلية مع التحرير السابق أو المسبق، فقد طبقته في صورة معدّلة بعض الشركات مثل شركة جنرال موتورز في كندا وشركة ماتيل للألعاب في امريكا وبعض الشركات اليابانية كذلك. فهذه الشركات تعطي تعليمات للذين يكتبون النصوص باللغة الاصلية بمراعاة قواعد معينة مثل: "لاتستعمل إلاهذه الكلمات. واستخدم هذه التراكيب. تجنب الجملة التي بها أكثر من فاعل واحد، ولاتستخدم المبني للمجهول...".من ثم يصبح النص المراد ترجمته "مفهوما" أو "مقبولاً" للحاسوب، فتأتي الترجمة سلسة وصحيحة لا تتطلب تنقيحا أو تحريراً يذكر.

غير أن هذا الاسلوب ممكن أحيانا لأن الشركات التي تتبع هذا الأسلوب في الترجمة الآلية هي التي تنتج كتيبات الإرشادات والكتبيات الإعلامية، أي النصوص المراد ترجمتها، وتستطيع أن تتحكم في تلك النصوص باعطاء التعليمات للمؤلفين.

وهناك أسلوب آخر للتحرير المسبق وهو أن نعطي النص الاصلي إلى محرر ليعيد صياغة النص بلغة يفهمها الحاسوب.ثم يعطى النص المعدّل للحاسوب ليترجمه. وكما هو واضح فهذا أسلوب عقيم جدا لأن الوقت الذي نضيعه في إعادة صياغة النص يمكن أن نقضيه في الترجمة مباشرة أي بالترجمة البشرية.

جـ - الترجمة التحاورية interactive:

وأما الأسلوب الثالث فنجده في اتجاه يسمى بالترجمة التحاورية، ويعرف أحيانا بالترجمة الآلية بمعاونة الإنسان. حيث نجد المترجم يجلس إلى الحاسوب وأمامه الشاشة أو المرقاب ولوحة الطباعة. وتتم الترجمة جملة جملة، ويظهر جزء من النص الأصلي مع ما يقابله من الترجمة. وإذا أشكل على الحاسوب شيء أو

لاحظ المترجم الذي يتابع عملية الترجمة على شاشة الحاسوب مشكلة ماتدخل بصور مختلفة، حسب نوع الإشكال الوارد في العملية .(وهذه أسهل طبعا من المراجعة الكاملة (أو التحرير اللاحق)، من وجهة نظر المراجع أو المحرر.) في هذه العملية قد يسأل الحاسوب مثلا "تربية هنا ماذا يقصد بها هل تترجم education, breeding or upbringing أم ماذا؟ والمترجم يقول له: ترجمها كذا. فيحفظ الحاسوب هذا. وفي المستقبل يستعمل هذه الصيغة. ولذلك سميت هذه بالترجمة التحاورية، حيث إن هناك تحاورا بين المترجم البشري وبين المترجم الآلي.

والمشكلة الأساسية في هذا الأسلوب من الترجمة الآلية ضرورة وجود المترجم أمام الحاسوب طوال عملية الترجمة. بينما في الأسلوبين الآخرين قد يقوم الحاسوب بالترجمة في خارج أوقات العمل الرسمية للموظفين، كما أن المحرر لايشترط وجوده في نفس المكان. من هنا قد تكون الترجمة هنا أبطأ منها في الأسلوبين الآخرين. ولا أذكر أن أيا من أنظمة الترجمة التحاورية له انتشار معروف في الوقت الراهن.

لمزيد من المعلومات حول أنظمة الترجمة المختلفة ومقارباتها approaches للترجمة الآلية، والتي لم يطرأ على مقارباتها تغيير يذكر منذ سنوات عديدة، ينظرصيني (١٩٩٦) وصالح (٢٠٠٠) والحميدان (١٩٩٨) و(١٩٩٩) Trujillo, Arturo .

١-١- الترجمة الآلية واللغة العربية:

١-١-١- المشروعات الأجنبية:

(١) يبدو أن محاولات الترجمة الآلية من الإنجليزية خاصة إلى العربية بدأت منذ أكثر من ثلاثة عقود من الزمان، حيث طور نظام سيستران Systran برنامجا للترجمة من الإنجليزية إلى اللغة العربية، ولكن لأسباب إدارية واقتصادية لم يكتب له الانتشار، والبرنامج المذكور يعمل على حاسوب كبير، ولعل هذا أحد أسباب عدم انتشاره، إضافة إلي الأسباب الأخرى. و لكن يبدو أن نظام سيستران لازال يعمل، ويمكن الاستفادة منه لقاء رسوم معلومة.

(٢) وقد تبعه نظام وايدنر Weidner الذي طور أيضا برنامج للترجمة من الإنجليزية إلى العربية، وكان يعمل على حاسوب متوسط وقد كان يستخدم في كل من الولايات المتحدة الأمريكية وفي الرياض لبعض من الوقت، ولكننا لانعلم ماحدث للبرنامج المذكور، خاصة وأن ملكية شركة وايدنر للاتصالات انتقلت إلى إحدى الشركات اليابانية منذ أكثر من عشر سنوات، كما أن الشركتين اللتين كانتا تستخدمان هذا النظام في

نيوجرسي بأمريكا(Language Management, Inc). وفي الرياض (شركة التعريب الآلي) توقفتا عن العمل في مجال الترجمة الآلية.

(٣) قامت منذ أكثر من عشر سنوات محاولتان تجريبيتان للترجمة الآلية من الإنجليزية إلي العربية، كانت الأولى في الولايات المتحدة الأمريكية قامت بها شركة ألبس Alps التي ما زال لديها برامج للترجمة بين عدد من اللغات، وتطبق مبدأ الترجمة التحاورية، ولكنها أوقفت العمل في تطوير برنامج الترجمة إلي العربية،منذ منتصف الثمانينيات، حسب ماذكره لي أحد المسؤولين عن الشركة المذكورة (ومقرها بروفو، يوتا) في اتصال هاتفي أجريته معه في عام ١٩٨٧م.

(٤) أما المحاولة الثانية فكانت في فرنسا لدى جامعة غرينوبل Grenoble، حيث كان الأستاذ الراحل فوكوا يعمل علي تطوير برنامج مبني على نظام أريان ٧٦ Arianne ٧٦، وذلك للترجمة من الإنجليزية إلي العربية، غير أن هذا المشروع توقف أيضا منذ أكثر من عشر سنوات، كما أنه كان يعمل على الحاسوب الكبير main frame.

١-١-٢- المشروعات العربية:

ربما كانت أول محاولة عربية لتطوير برنامج للترجمة من الإنجليزية إلى العربية محاولة الدكتور بشاي الأستاذ السابق بجامعة هارفارد، وذلك منذ أوائل السبعينيات. كان البرنامج يطمح إلي تطوير نظام عامل مبني علي منهج التحرير السابق، أي تحرير النص الإنجليزي قبل ترجمته آليا إلى اللغة العربية. ونظرا لتكلفة هذا التحرير السابق من حيث الجهد والمال والوقت لم يلق البرنامج المذكور قبولا يشجعه علي الاستمرار كما يبدو.

ولكن منذ أواخر العقد الماضي وخلال العقد الميلادي الحالي قامت عدة محاولات عربية أخرى لتطوير أنظمة للترجمة إلى اللغة العربية، بعضها لازال قيد التطوير، كما هو الحال مع برنامج ترجمان التونسي والبرامج الأخرى التي تعمل عليها عدة جهات في مصر والأردن. ثم هناك الأنظمة العاملة والمتوافرة حاليا في الأسواق، وهناك عدة أنظمة معروفة لدينا، كلها تعمل على الحاسوب الشخصي وهي:

أولا: أنظمة مستقلة stand alone

(١) نظام "المترجم العربي" الذي طورته شركة ATA في لندن، ولدى الشركة المذكورة فرع في مسقط بعمان. وقد طورت الشركة المذكورة برنامجا مصغرا أسمته "الوافي ". ويترجم من الإنجليزية إلى العربية. وهو من المشروعات العربية الرائدة في ميدان الترجمة الآلية.

(٢) نظام "الناقل العربي" الذي طورته شركة سيموس Cimos العربية في باريس، وهذا النظام أكثر الأنظمة طموحا، حيث لدى الشركة المذكورة أربعة برامج للترجمة بين الإنجليزية والعربية وبين الفرنسية والعربية - برنامج لكل اتجاه. وتذكرالشركة أن النظام يستخدم "قاعدة النصوص المترجمة سلفا" أي أنه يستفيد من فكرة ذاكرة الترجمة التي نتحدث عنها أدناه.

وهناك في الأسواق برنامج يدعى "المترجم المحترف"، يبدو أن له علاقة بالناقل، كان يباع مع مجموعة من المعاجم المتخصصة.

(٣) نظام الترجمة الألية من إنتاج شركة صخر، يدعى(Sakhr Enterprise Translation –SET). و-يعمل في اتجاهين (عربي – إنجليزي وبالعكس).كما تذكر الشركة بأن النظام يتيح أيضا نظام ذاكرة الترجمة.

ثانيا: أنظمة ترجمة على الشابكة internet

(١) المسبار، وهو من إنتاج شركة ATA، منتجة "المترجم العربي".

(٤) نظام الترجمة الآلية على محرك Google، وهو في اتجاهين: عربي-إنجليزي وإنجليزي-عربي.

(٣) نظام ترجم (صخر) Tarjim، ويعمل في اتجاهين كذلك.

(٤) البرنامج التابع لشركة سيموس (الناقل) في اتجاهين: عربي-إنجليزي/فرنسي وبالعكس.

(٥) نظام بابيلون Babylon ويبدو أنه النظام ا لمستخدم في برنامج مايكروسوفت وورد لتنسيق النصوص Microsoft Word

ثالثا: أنظمة للترجمة الشفوية على الهواتف الذكية:

.Sakhr Mobile S٢S Arabic Translator for Government & Enterprise (for iPhone)

- برامج iTranslate, Translate, SayHi, Voice وغيرها. وجديربالذكر أن بعض التطبيقات على الهواتف الذكية، مثل Triplens وغيرها، تتيح التقاط صورة النص المكتوب وترجمته شفويًا أو مكتوبًا إلى اللغة الهدف.

٢. معيينات حاسوبية للترجمة:

لاتقتصر استخدامات الحاسوب على الترجمة الآلية مباشرة، أي ترجمة النصوص الكاملة بوساطة الحاسوب. بل هناك معينات حاسوبية أخرى في الترجمة مثل ذاكرة الترجمة، المعاجم الإلكترونية.

١-٢-١- ذاكرات الترجمة:

يخلط بعض الباحثين بين مفهومي الترجمة الآلية (أو بمعاونة الحاسوب) وذاكرات الترجمة. ونود أن نشير أن هناك فرقا جوهريا بين النظامين. ففي الترجمة الآلية بصورها المختلفة لا يفترض النظام توافر نصوص مترجمة ثنائية اللغة لينقل منها. بل نقدم النص المراد ترجمته ليقوم الحاسوب بالترجمة ابتداء، ثم نجري ما نشاء من تحريرعلى النص المترجم. أما فيما يسمى بذاكرات الترجمة translation memories، فلا بد من توافر نصوص مترجمة ومخزنة في ذاكرة البرنامج. وكل ما يقوم به البرنامج هو المطابقة بين النص الجديد وما هو مخزن في ذاكرة الحاسوب وإيراد الترجمة أو الترجمات المتاحة. من ثم كلما زاد عدد النصوص ثنائية اللغة المخزنة كان النتيجة أفضل. وجدير أن هذه الأنظمة تتطلب جهدا من المترجم أو المترجمين لإجراء المواءمات بين النصوص الأصلية وترجماتها قبل الاستفادة منها في عملية الترجمة. (انظر المقارنة أدناه.)

هذا وينصح بمثل هذا البرنامج في المؤسسات التي لديها الكثير من النصوص المتشابهة التي تتطلب الترجمة. كما يلاحظ أن هذه الأنظمة بها برامج فرعية لمعالجة المصطلحات وإعداد قوائم بها.

ومن أشهر الأنظمة متعددة اللغات:

Déjà Vu, Wordfast Pro, Trados

١-٢--١-١- مقارنة بين الترجمة الآلية وذاكرة الترجمة:

حيث إن أشيع طريقتين في استخدام الحاسوب في الترجمة هما الترجمة الآلية machine translation وذاكرة الترجمة translation memory، سنحاول فيما يلي إلقاء الضوء على بعض الاختلافات المهمة بينهما:

- عام وخاص
- جاهز للاستعمال أو يحتاج إلى تهيئة كثيرة.

- دقة الترجمة وسلاسة الأسلوب.

- من حيث العمومية والخصوصية:

يعتبر نظام الترجمة الألية نظاما عاما يستخدمه الجمهور لأغراض مختلفة، بينما نجد أن فائدة ذاكرة الترجمة ترتبط بنوعية النصوص التي نحتاج إلى ترجمتها، حيث إنها نوع من قاعدة المعطيات للنصوص المترجمة، من ثم فإن فائدتها أكير للأفراد أو الجهات التي يتسم عملها بشيء من التشابه والتكرار.

- ن حيث جهد التهيئة:

من الواضح أن الذين جربوا نظام الترجمة الآلية (مثل المتاح على google) أدركوا أنهم لا يحتاجون إلى تهيئة النظام للاستفادة منه، بل إنهم لا يستطيعون ذلك وإن أردوا. ولكن فاعلية النظام تكون أفضل ولا شك لو كان يتيح للمسخدم إجراء بعض التعديلات، من مثل إضافة كلمات أو مصطلحات جديدة مع مقابلاتها، أوإذا كان النظام "ذكيا" يتعلم من تجاربه السابقة. وهذه الإمكانات ليست متاحة إلا في أنظمة الترجمة الخاصة stand alone، مثل المترجم العربي والناقل العربي وغيرهما، أوالتي تشرف عليها مؤسسات تقدم خدماتها برسوم للمستفيدين، مثل Systran.

أما ذاكرة الترجمة، فهي بطبيعتها تتطلب تهيئة من قبل مستخدمها ، ويشمل ذلك تخزين النصوص المتوافرة لدى الشخص أو الجهة المستفيدة وترجماتها، ثم تقطيع تلك النصوص segmentation وإجراء المواءمة بين المقاطع المختلفة في كل من النص وترجمته، وهو مايسمى بالمحاذاة alignment. فبدون هذا لن يعمل برنامج ذاكرة الترجمة. ولاشك أن كفاءة النظام تعتمد على جودة مثل هذه العمليات وعلى كمية النصوص المترجمة المخزنة في هذه الذاكرة.

- من حيث النصوص المترجمة:

لاشك أن النصوص المترجمة عن طريق ذاكرة الترجمة أعلى جودة وأفضل أسلوبا، بشرط مراعاة ذلك في النصوص المخزنة وترجماتها. أما نظام الترجمة الآلية فيخضع منتجه على جودة النظام وبرمجته التي لا تعد غالبا لمستخدم معين أو نصوص محددة، من ثم نجد العيوب والأخطاء التي نتحدث عنها لاحقا عند مناقشة أخطاء الترجمة الآلية.

في ضوء ماسبق، نستطيع القول بأن نظام الترجمة الآلية الذي يستفيد من ذاكرة الترجمة، يتوقع أن تكون مستوى الترجمة لديها أعلى من غيرها.

١-٢-٢- قواعد معطيات النصوص المترجمة:

تشبه هذه إلى حد ما ذاكرات الترجمة ولكنها لاتقوم بالترجمة، بل توفر للمترجم نماذج من النصوص المترجمة ليستفيد منها في عمله. ويمكن لأي مترجم ملم بالبرمجيات الحاسوبية إعداد مثل هذه القواعد والاستفادة منها.

وجدير بالذكر أن الشابكة (الإنترنت) كنز غير محدود للنصوص المترجمة أو المكتوبة باللغة الهدف (التي نرغب في الترجمة إليها)، مما يعين المترجم على ترجمة النصوص أو الفقرات أو العبارات المختلفة أو للتثبت من صحة ترجمته لبعض النصوص أو العبارات. وفي واقع الأمر نجد أن بعض القوانين العربية هي مجرد ترجمات لقوانين أجنبية، مع تعديلات قليلة أحيانا، وقد يعثر المترجم على تلك الأصول.

٣,١. المعينات المعجمية في الترجمة:

هناك عدد من الأساليب التي يعين فيها الحاسوب في البحث عن الألفاظ (بل والتعبيرات) المقابلة لألفاظ أو تعبيرات يود ترجمتها، ومن أقدمها بنوك المصطلحات الآلية.

١-٣-١ بنوك المصطلحات الآلية:

من الاستعمالات المبكرة نسبيا للحاسوب في خدمة الترجمة بنوك المصطلحات الآلية (يسميها البعض بنوك المعطيات المصطلحية terminology data banks). وهناك عدد من بنوك المصطلحات الآلية في العالم (نذكر منها مثلا لكسيس Lexis في وزارة الدفاع الالمانية في بون ، وتيم TEAM في ميونيخ (تابع لشركة سيمنز)، ويورديكاوتوم Eurodicautom في لوكسمبورغ (يتبع أمانة الاتحاد الأوربي)، وتيرميوم Termium في كندا (يتبع الحكومة الكندية) و دانتيرم Danterm الدانمركي ونورماتيرم Normaterm الفرنسي و"باسم" في الرياض، و"معربي " في الرباط. كما أن هناك أعداداً كبيرة من هذه البنوك أنشئت في مناطق مختلفة من العالم في الآونة الأخيرة، بعضها حكومي وبعضها تجاري. ومن البنوك المتخصصة الحديثة نسبيا بنك المصطلحات الخاص بالأمم المتحدة UNTERMK ويقدم عبارات

كاملة (مثل معاهدة كذا وكذا) في لغات الأمم المتحدة الست.

يجدر بالذكر أن بنك المصطلحات ليس مجرد معجم محوسب، بل هو قاعدة معطيات يشمل المقابلات المصطلحية في عدد من اللغات مع مجموعة من البيانات الإضافية، مثل التعريفات والمعلومات التوثيقية، مثل مصدر كل مصطلح ودرجة توثيقيته وتصنيفه إلى غير ذلك. (انظر Sieny, ١٩٨٦). كذلك يمكن للباحث أن يبحث عن معلومات غير المقابلات المصطلحية.

يساعد بنك المصطلحات المترجم في تقديم المصطلح حسب الطلب.مثلا إذا أردت أن تعرف معنى كلمةneutron و مقابلاتها بالعربية أو الفرنسية اوبأية لغة أخرى يتعامل معها البنك المعني، تكتب الكلمة،أو تدخلها في الحاسوب بالأصح، فتظهر لك على الشاشة معنى الكلمة المقابلة في اللغة التي تريدها من ذاكرة الحاسوب. فيغنيك بذلك عن البحث في عشرات المعاجم ومصادر المصطلحات الأخرى، ممايوفر كثيرا من الجهد والوقت والمال كذلك. وهذا أبسط استخدام لاستخداما ت بنوك المصطلحات.

وبعض بنوك المصطلحات أكثر تطورا من ذلك، حيث يمكننا أن نقدم النص المراد ترجمته للحاسوب فينظر فيه، ويعطيك مقابل أية كلمة موجودة في ذاكرته . ومن ثم يظهر لك النص مطبوعا بعد إضافة ترجمات المصطلحات (مثلا الكلمات الانجليزية ومقابلاتها العربية او الكلمات الفرنسية ومقابلاتها العربية) إلى أن ينتهي من النص كله.

وهناك أسلوب آخر لعمل بنك المصطلحات، وهو أن ندخل النص الأصلي في الحاسوب، فيقارن الحاسوب النص بما في ذاكرته ثم يعطيك جميع المصطلحات التي وردت في النص مع مقابلاتها، مرتبة حسب ورودها في النص، أي أن الكلمة التي وردت في السطر الأول تسبق التي وردت في السطر الثاني أو الثالث وهكذا.

ثم هناك أسلوب ثالث وهوأن ينظر الحاسوب في النص ويقارن بما في ذاكرته من مصطلحات ويعطيك الكلمات مع مقابلاتها في قائمة مرتبة ألفبائيا، وهكذا يوفر للمترجم معجما خاصا به، ويوفر عليه بذلك

عملية البحث عن الكلمات في القواميس المختلفة.

أما الأسلوب الرابع للاستفادة من بنك المصطلحات فهو أن نطلب من الحاسوب أن يعطينا مثلا جميع المصطلحات المتوفرة في حقل معين من حقول التخصص مع مقابلاتها. فيتوفر لدينا معجم متخصص نرجع إليه وقت الحاجة. هذا وقد يكون المعجم في صورة أوراق مطبوعة أو قرص مضغوط CD.

١-٣-٢-بنوك المصطلحات الآلية في العالم العربي:

تبين لنا الدراسة التي أجريناها في وقت سابق أن هناك أربعة بنوك آلية للمصطلحات في العالم العربي، هي:

(١) مُعربي التابع لمركز دراسات الترجمة والتعريب في الرباط وهو أقدم بنوك المصطلحات في العالم العربي، وكان يطلق عليه لكسار سابقا.

(٢) باسم BASM التابع لمدينة الملك عبد العزيز للعلوم والتقنية KACST في الرياض، ويرجع تاريخ إنشائه إلي حوالي عام ١٩٨٣م.

(٣) بنك المصطلحات الموحدة التابع لمكتب تنسيق التعريب في الوطن العربي، والذي يجمع كل مسارد المصطلحات التي أنتجها المكتب خلال عدة عقود باللغات العربية والإنجليزية والفرنسية. وعنوان المكتب هو www.Arabization.org.ma

(٤) لعل من أحدث بنوك المصطلحات وأفضلها في مجال مصطلحات التقنية هو أرابتيرم Arabterm الذي أنشئ بالتعاون بين مكتب تنسيق التعريب وأحد أجهزة الحكومة الألمانية للتعاون الدولي. ويشتمل البنك على المصطلحات التقنية بأربع لغات (العربية والإنجليزية والفرنسية والألمانية)، مع شرح باللغة العربية، كما يشتمل الشرح أحيانًا على رسومات توضيحية.

(٦) إضافة إلى ماسبق يمكننا أن نذكر بنوك أو معاجم المصطلحات المتخصصة لدى الأمم المتحدة ومنظمة الصحة العالمية (المعجم الطبي الموحد)، وكلها تتعامل مع المصطلحات العربية ذات العلاقة بالمنظمة.

وجدير بالذكر أن لبنوك المصطلحات الآلية العربية فوائد هامة أخرى في خدمة التعريب، ألا وهو توثيق المصطلحات ونشرها وتنسيق الجهود في مجال العمل المصطلحي.

(انظر الفصل الخاص ببنوك المصطلحات في هذا الكتاب.)

١-٣--٣- المعاجم الإلكترونية والمحوسبة:

من المعينات الواسعة الانتشار للترجمة المعاجم المحوسبة أحادية اللغة وثنائيتها ومتعددة اللغات. التي تخدم المترجم في فهم النصوص وفي ترجمتها. وهي تقدم المقابلات اللفظية بين اللغات المخزونة في ذاكرة الحاسوب. وتتوافر في الأسواق أعداد من هذه المعاجم في صورة أقراص للاستعمال علي الحواسيب الشخصية وهي في تطور مستمر من حيث النوع وحجم مخزوناتها. فبعضها لايكتفي بتقديم الألفاظ ومقابلاتها بل تقدم كذلك بعض التعبيرات والجمل الشائعة ومقابلاتها في اللغة الأجنبية.

وجدير بالذكر أن بعض برامج تنسيق النصوص متعددة اللغات أو ثنائيتها توفر معجما آليا ثنائي اللغة لمستخدميها، إضافة إلى التدقيق الإملائي.

وينبغي أن نشير إلى أن هناك أعدادا كبيرة من المعاجم المتاحة على الإنترنت، وفي الهواتف الذكية كذلك، بعضها إلكتروني صمم للاستعمال الحاسوبي (التفاعلي) وبعضها نسخ محوسبة لمعاجم مطبوعة.

١-٣-٢-١-المعاجم أحادية اللغة:

لابد لنا من الإشارة هنا إلى أن المعاجم أحادية اللغة قد تكون أداة لفهم النصوص أو للتعبير بلغة ما. وليس المقام مناسبا للخوض في تفاصيل ذلك. ولكن يكفينا هنا التمييز بين معاجم الألفاظ والتعبيرات الاصطلاحية والأمثال من جهة، وكلها أدوات معينة في فهم النصوص المسموعة والمقروءة.

وهناك عدد من معاجم الألفاظ العربية والأجنبية المحوسبة التي يستطيع المترجم الاستفادة منها. بل إن بعض برامج تنسيق النصوص تتوافر بها مكانز(معاجم مترادفات).

١-٣-٢-٢- المكنز (معجم المترادفات) الآلي:

يتوافر مع كثير من برامج تنسيق النصوص المكنز thesaurus الآلي حيث يقدم للكاتب والمترجم المرادفات ليختار الأصح أو الأنسب منها. كما أن هناك معاجم مستقلة لهذا الغرض.

١-٤- معينات حاسوبية أخرى:

١-٤-١- منسق النصوص:

اما الأسلوب الرابع من أساليب الاستفادة من الحاسوب فهو الاستعانة بمنسق النصوص أو معالج الكلمات

word processor، وهو نظام حاسوبي يكتب به المترجم النصوص المترجمة، بدلا من الآلة الكاتبة (الراقنة) سابقًا. ومن حسناته أنه يتيح لنا أن نعدل في النص بسهولة ويسر. ومن المعلوم أنه باستخدام منسق النصوص يمكننا أن نضيف كلمات أو أسطرأو فقرات كاملة أونحذفها وأن نقدم أونؤخرما نشاء بطريقة سهلة ميسورة جدا. هذا بالاضافة إلى أنه يمكننا استخدام منسق النصوص أحيانا بطريقة توفر عملية صف الحروف ؛ فبعض برامج تنسيق النصوص تسمح لنا بأن نأخذ الترجمة من الحاسوب إلى المطبعة مباشرة. وذلك بالإضافة إلى أنظمة "النشر المكتبي" التي تتيح للمترجم تجهيز النص المترجم للنشر مباشرة، وذلك بطبع الترجمة بطابعات الليزر، دون الحاجة إلى دور الطباعة، إذا كانت عدد النسخ محدودة.

إضافة إلى ماذكرنا نجد أن عددا من برامج تنسيق النصوص توفر إمكانات أخرى مفيدة للكاتب والمترجم، من أهما مايلي:

- التدقيق الإملائي.

- التدقيق النحوي والأسلوبي.

١-٤--٢- برمجيات قواعد البيانات، مثل Access والحسابات spread sheet، مثل Exel وتفيد هذه المؤسسات والأفراد في إعداد معاجم أو مسارد glossaries خاصة بالمؤسسة أو شخصية تتعلق بالمترجم، مما يحتاج إلى الرجوع إليه بصورة متكررة. وتتيح هذه البرمجيات إمكانية ترتيب المصطلحات ومقابلاتها في أي من الاتجاهين (اللغة الأولى أوالثانية)، إضافة إلى تصنيف المصطلحات حسب الموضوعات، مع إضافة أية معلومات أخرى، مثل التعريف أو المصدر... مما يتيح لنا بناء بنوك مصطلحات خاصة بالفرد أو المؤسسة.

١-٤--٣- نظام الإملاء الحاسوبي:

حيث إن بعض المترجمين المتمرسين لا يحسنون الكتابة على الحاسوب بالسرعة المطلوبة، فهم يلجؤون إلى الإملاء، حيث يقوم ناسخ محترف بكتابة النص من النص الشفوي. ونجد أن الحاسوب يقدم خدمة ممتازة لهذه الفئة من المترجمين في صورتين:

(أ) برنامج التسجيل الصوتي الحاسوبي: وهو نظام مطور لما كان يعرف بالدكتافون dictaphone الذي يستخدمه بعض المديرين والمترجمين المحترفين في تسجيل ترجمتهم صوتيا، ليتم تحويله في وقت لاحق إلى نص مكتوب أو مطبوع. أما دور الحاسوب في التطوير فهو أنه يسمح للمترجم

أن يجري التعديلات صوتيا على جهاز الإملاء، من حذف وإضافة وتعديل دون تأثير على بقية النص، تماما كما يتم التعامل مع منسق النصوص المكتوبة المذكور أعلاه. إضافة إلى ذلك يسمح البرنامج المذكور للمترجم أن يسجل النص في حاسوب موجود على بعد مئات أوآلاف الأميال من مكانه. وهذه إمكانات لم تكن متاحة مع أجهزة التسجيل الصوتي التقليدية.

(ب) الإملاء الآلي: وهو برنامج حاسوبي يحول اللغة المنطوقة مباشرة إلى لغة مكتوبة، مثل برنامج Dragon الذي يعمل مع اللغة الإنجليزية. ولابد لنا هنا من التنويه أن ذلك يتطلب أن يكون نطق المملي جيدا وصحيحا حتى يقوم البرنامج المذكور بعمله بصورة جيدة. لذلك نجد أن البرنامج المذكور له صورتان أساسيتان: إحداهما للإنجليزية الأمريكية وأخرى للبريطانية. إضافة إلى ذلك يتيح البرنامج إمكانية التكيف مع لغة المستخدم، حيث يتم "تدريب" البرنامج على نطق المستخدم وصوته ليتعرف عليهما مستقبلا. هناك حاليا إمكانات للإملاء الآلي للغة العربية، بعصها متاح مع منسق النصوص وعلى بعص الهواتف الذكية.

١-٤-٤- البريد الإلكتروني:

تلجأ كثير من المؤسسات الكبيرة إلى نظام "المترجم المستقل freelance " الذي تتعاقد معه طبقا للاحتياجات، بدلا من توظيف عدد كبيرمن المترجمين لديها بصورة دائمة. كما تعمد بعضها إلى الاستعانة بمكاتب الترجمة. في مثل هذه الحالات ترسل إلى المترجم أو مكتب الترجمة المواد المراد ترجمتها، وتعاد بعد الترجمة إلى المؤسسة. هنا نجد أن من الوسائل الناجعة لتبادل المعلومات وأكثرها اقتصادية نظام البريد الإلكتروني، وهو أمر أضحى أكثر شيوعا مع انتشار استخدام شبكة الإنترنت.

١-٤-٥- الشابكة:

لا شك أن الشابكة تمثل مصدرًا مهمًا للمترجم، ليس للترجمة الآلية فحسب، بل وللعثور على نماذج للترجمات أو النصوص الأصلية تشبه ما يترجم إليه من نصوص، إضافة إلى التثبت من صحة استخدام بعض العبارات في اللغة الهدف حسب سياقاتها أو تخصصاتها.

٢- تقويم أنظمة الترجمة الآلية:

يتحدث أرتورو تروخييو Arturo Trujillo في كتابه عن الترجمة الآلية عن فئات المعنيين بتقويم الترجمة الآلية، وهم الباحثون وممولو الأبحاث والمطورون والمشترون والمترجمون والمستقبلون (للنص المترجم). أما ما يهم المترجم فيشمل:

- الجوانب التشغيلية والتفاعلية مع نظام الترجمة وسهولة الاستخدام،
- إمكانات التنقيح المسبقة واللاحقة.
- نوع الأخطاء اللغوية. ويذكر مثالا على ذلك ترجمة قائمة المصطلحات بين الإنجليزية والأسبانية، فإذا كان نظام الترجمة مثلا يغفل إضافة أداة التعريف للمصطلحات الأسبانية فإنه سيكلف المترجم جهدا كبيرا لإضافة أداة التعريف المناسب لكل مصطلح (جدير بالذكر أن الأسبانية تتطلب مطابقة بين أداة التعريف والاسم من حيث التذكير والتأنيث والإفراد والجمع).
- سهولة إضافة المصطلحات واستقبالها وتصديرها.
- إمكانية قراءة النصوص بصيغ formats مختلفة.
- توافر الخدمات المساندة.
- إمكانية الاتصال بمستخدمين آخرين للنظام. (Trujillo, ١٩٩٩: ٢٥١-٢٥٦).
- ويمكننا أن نضيف الجهد المطلوب لتهيئة النظام للعمل والاستفادة منه. فالترجمة الآلية مثلا لا تتطلب جهدا خاصا للاستفادة منه مباشرة، ولو أنه قد يحتاج إلى تحديث أوزيادة في معجمه العام أو الخاص. أما ذاكرة الترجمة فتتطلب جهدا كبيرا قبل الاستفادة منه، كما نبين في موضع آخر من هذه الورقة.

٢-١- تقويم الترجمة:

يذكر الباحث المذكور ثلاثة أساليب لتقويم الترجمة، هي:

٢-١-١- سهولة الفهم أوالمفهومية intelligibility :

ويقصد بها وضوح النص المترجم من حيث وضوح اللغة وسلاسة الأسلوب، بغض النظرعن دقته في نقل معاني النص الأصل. وهناك عدة طرق لعمل ذلك.

(أ) التقويم الإجمالي: واقترح بعضهم سلما من أربع درجات لتقويم ذلك، بينما اقترح آخرون عشر درجات لذلك، تبدأ من "غير مفهوم أبدا" إلى "مفهوم تماما". أما السلم الذي يستخدم كثيرا فهو السلم الذي اقترحه ناغاو وآخرون (١٩٨٨) .Nagao et al والذي يتكون من خمس درجات:

"١. معنى الجملة واضح ، وليست هناك تساؤلات. كما أن النحو واستعمال الألفاظ و/أو الأسلوب كلها مناسبة، ولا تحتاج إلى إعادة صياغة.

٢. معنى الجملة واضح، وليس هناك تساؤلات، ولكن هناك مشكلات في النحو واستعمال الألفاظ و/أو الأسلوب كله مناسب، مما يجعل المستوى الإجمالي أقل من ١.

٣. المعنى الأساسي للجملة واضح ، ولكنك لست متأكدا من بعض التفاصيل نتيجة لوجود مشكلات في النحو واستعمال الألفاظ. وستحتاج إلى أن تنظر في الأصل لاستيضاح المعنى.

٤. الجملة بها الكثير من المشكلات النحوية وفي استعمال الألفاظ، ولا يمكنك سوى تخمين المعنى بعد دراسة كافية للنص.

٥. لايمكننا فهم الجملة على الإطلاق.

حيث أن السلم المذكور يفتقر إلى الموضوعية، فما يصنفه أحدهم ١ قد يصنفه آخر ٢. وللتقليل من عامل الذاتية يقترح وجود أكثر من مقوم، وتجمع الدرجات التي يعطيها هؤلاء المقومون ويستخرج متوسط الدرجات.

(ب) هناك أسلوب آخر لتقويم سهولة الفهم، وهو اختبار التتمة cloze test، حيث تحذف كلمات من النص بصورة منتظمة (مثلا كل خامس كلمة)، ويطلب من بعض القراء الاتيان بالكلمات المناسبة لملء هذه الفراغات. ثم ننظر في مدى الارتباط بين الكلمات المحذوفة والكلمات المقترحة من حيث قسم الكلام والمعنى، لنستدل من ذلك على مفهومية النص.

(جـ) وهناك إمكانية اختبارات الاستيعاب comprehension، حيث يعطى النص لبعضهم ويجيبون بعد قراءته على أسئلة الاستيعاب المعروفة <مثل الاختيار من متعدد أوصحيح /خطأ/ لاأدري...>" (Trujillo, ١٩٩٩: ٢٥٨-٩).

٢-١-٢- الدقة accuracy:

يقول تروخيو:"بعد قياس مفهومية النص نقوم بقياس دقة الترجمة... فالجملة المفهومة تماما قد لا تعكس المعنى الصحيح للنص الأصل نتيجة لخطأ في تفسيره. وهناك عدة وسائل لذلك، منها:

(أ) تقديم مجموعة من الجمل في نص أصل مع مقابلاتها للمقومين، مع سلم لتقويم دقة الترجمة. ويقترح ناغاو وآخرون (١٩٨٨) السلم التالي ذي السبع درجات لتقويم الدقة، نعرضها أدناه مع شيء من التعديل:

١. تم نقل معنى الجملة الأصل بأمانة، كما أن الجملة المترجمة واضحة للناطق الأصلي باللغة الهدف، وليس هناك حاجة لإعادة الكتابة.
٢. تم نقل فحوى الجملة الأصل بامانة إلى اللغة الهدف، والترجمة مفهومة للمتحدث بها، ولكن هناك حاجة إلى المراجعة.
٣. تم نقل فحوى الجملة الأصل بأمانة في اللغة الهدف، ولكن هناك حاجة إلى تغييرات على مستوى النحو (تركيب الجمل).
٤. على الرغم من أن جملة النص الأصل تم نقلها بأمانة بشكل عام في اللغة الهدف، هناك مشكلات مثل العلاقة بين التعبيرات والتعابير phrases and expressions وبالزمن النحوي وصيغ الجمع ومواقع الظروف. كما أن هناك تكرار لبعض الأسماء في الجملة.
٥. لم يتم نقل فحوى الجملة الأصل بصورة مناسبة في اللغة الهدف. فهناك عبارات ناقصة، كما أن هناك مشكلات في العلاقات بين مكونات الجملة الرئيسة clauses وبين هذه والعبارات phrases أو بين عناصر الجملة.
٦. لم يتم نقل فحوى الجملة الأصل.
٧. لم يتم نقل فحوى الجملة الأصل على الإطلاق. كما أن النتيجة ليست جملة تامة، فالمسند أو المسند إليه ناقص...إلى غير ذلك.

مرة أخرى، نجد أن هذا السلم لا يمكن تطبيقه بصورة موضوعية، حيث إن ذلك يعتمد على مدى صرامة المقومين.

(ب) هناك أسلوب آخر في قياس دقة الترجمة. ويكمن في تقويم الأداء العملي لترجمة الكتب الإرشادية، حيث ننظر في مدى قدرة قارئ الترجمة على تطبيق التعليمات عمليا....

(جـ) كما أن هناك طريقة أصعب من السابقتين، وهي الترجمة العكسية، أي إعادة ترجمة النص المترجم إلى اللغة الأصل، ومقارنة النتيجة بالنص الأصل.

٢-١-٣- تحليل الأخطاء error analysis

هناك أسلوبان رئيسيان في جمع المادة لتحليل الأخطاء، هما المادة التلقائية spontaneous data والمادة المستنطقة elicited data. ونجد في الدراسات العربية الخاصة بالترجمة الآلية تطبيقا لهذين الأسلوبين.

٢-١-٣-١- أسلوب المادة الاختبارية:

وهوالمنهج الذي اتبعه داود عبده (وهو لساني عربي معروف)، حيث استخلص الباحث الأخطاء عن طريق اختبار ثلاثة أنظمة ترجمة آلية بترجمة بعض العبارات والجمل التي توقع حدوث أخطاء فيها في ضوء مقارنة بين العربية والإنجليزية. ونورد ملخصا لدراسة داود عبده أدناه.

أهم أخطاء الترجمة الآلية عند داود عبده:

١- عدم اكتشاف العلاقة بين كلمتين:

أ- هي طالب

ب- أختي ليست طالبا

ت- هذة البنات لسن مدرسين. هم طلاب.

ث- هؤلاء البنات لسن معلمات. هم طلاب

ج- عمر وسامي طالب.

ح- عمر وسامي...هم يدرسون.

خ- هذا هو الطالب تقريبا من تكلمت This is the student about whom I talked

د- سيارتك أفضل من المنجم mine

ذ- هل قرأت هذا الكتاب؟ نعم قد / نعم عندي

ر- ساذهب إذا أنت س.

ز- قابلت أخاه، الأب والأخت.

س- فرأت كتابا الذي اشتريت أمس.

ش- نسي هذا طالب كتابه.

ص- هذا صديق لملكك ليس طالبا This friend of yours

ض- هذا أخذ مكانا (حدث) took place

ط- ماذا على الأرض ?...What on earth

ظ- ضع الضوء على .Turn the light on

ع- كان معلم، أكلت كثير، خمس دقائقا.

غ- جاء الطالبان الذين... الطالباتت الذين...

٢- المطابقة بين المبتدأ والخبر في التذكير والتأنيث، كما في الأمثلة (أ) و (ب) و(ت) أعلاه.

٣- عدم ملاحظة العلاقة بين الجمل (تعتبر كل جملة منفصلة تماما عن سابقتها) كما في المثالين (ت) و(ث).
٤- عدم التنبه إلى الأخطاء الإملائية الواضحة (Omar and Sami are student حيث ترجمت "عمر وسامي طالب".
٥- معاملة المثنى جمعا، كما في المثالين (ج) و(غ).
٦- الترجمة الحرفية أحيانا للمسكوكات (العبارات الاصطلاحية)، كما في الأمثلة (خ) و (ض و(ط) و(ظ).
٧- الترجمة الحرفية مخالفا لقواعد العربية، كما في المثال (ص).
٨- مشكلة السياق اللفظي، كما في المثال (د).
٩- مشكلة العبارات المحذوفة لفظا، مع وجودها معنى، كما في الأمثلة (ذ) و (ر) (ز).
١٠- وجوب حذف الاسم الموصول في العربية بعد النكرة، كما في المثال (س).
١١- إضافة أداة التعريف قبل المشار إليه، كما في المثال(ش).
١٢- أخطاء إعرابية ، كما في المثالين (ع) و(غ).
١٣- المطابقة بين الاسم والاسم الموصول الذي يرتبط به، خصوصا في التثنية والتذكير والتأنيث.

٢-١-٣-٢- تحليل المادة التلقائية:

أما الأسلوب الثاني فهو ما اتبعه الدكتور إلهام عفيفي الذي قام فيه بتحليل الأخطاء التي وردت في ترجمة نص كامل من الإنجليزية إلى العربية حاسوبيا، أي أنه اعتمد على المادة المترجمة كاملة بصورة آلية.

أهم الأخطاء عند إلهامي عفيفي:

١- الخطأ في الحركات الإعرابية ("... تمثل تجمع ضخم، أن هناك أنواع..."
٢- المطابقة بين الفاعل والفعل ("المرحلتان الرابعة والخامسة تهدف إلى..." "مهما اختلف مستويات..."، "كما أن استخدام عينتين عشوائيتين يؤديان...") الأخير قد يكون له مايفسره.
٣- إغفال تفضيل العربية للبدء بالفعل ("نماذج التقويم تصف...")
٤- استخدام عبارات وصفية غير مستساغة بدلا من عبارات تتضمن شبه الجملة ("ندوات تقويمية" بدلا من "ندوات عن التقويم".
٥- "ترجمة عبارات ترجمة حرفية خطية" ("محك الإتقان التحصيل المقرر" بدلا من "محك إتقان التحصيل في المقرر" <زيادة ال التعريف وحذف حرف الجر>.
٦- "استخدام ألفاظ أو عبارات غير مستساغة، مثل ("رجال العمل الاجتماعي" بدلا من "الأخصائيون الاجتماعيون"

٧- تعريب الكلمات الأجنبية التي لا مقابل لها في معجم البرنامج، ويعني ذلك كتابة نطق الكلمة الأجنبية بالحرف العربي.

٨- عدم الاطراد في استخدام المقابلات المترادفة أو شبه المترادفة ("اختبار / قياس) لكل من test, measurement).

٩- "عدم إضافة لفظ لم يرد في جملة من جمل النص الأصل وإنما يفهم من سياق النص ولا تستقيم الترجمة بدونه" ("تلك تستحق التسجيل" بدلا من "تلك الظاهرة تستحق التسجيل").

١٠- "أما عن قصور الترجمة الآلية من حيث الأسلوب، فإنها كثيرا ما تأتي بجمل وفقرات ضعيفة في لغتها وصياغتها، فتبدو للقارىء كلمات مرصوصة لا رابط بينها..." (عفيفي: ٣٩٨)

١١- مشكلة المصطلحات والمختصرات ("IRT, SEP).

١٢- "القصور الشكلي في الترجمة الآلية:

١٣- "عدم الربط بين الجمل والفقرات في النص العربي (مثل و، ف، ثم، ومن ثم...) لاكتفاء الإنجليزية بعلامة الوقف – النقطة- دون الحاجة إلى أداة ربط في الجملة التالية.

١٤- عدم ملاحظة معاني بعض علامات الترقيم في الإنجليزية (مثل الفاصلة عند التعداد والتي يجب أن تترجم واوا، استخدام الواصلة (-) في الإنجليزية أحيانا بين العبارات الوصفية المركبة، في حين أن اللغة العربية تعبر عن مثل هذه التركيبات بمضاف ومضاف إليه." (٤٠٠).

١٥- عدم ملاحظة استخدام الإنجليزية لبعض الأبناط لعناوين الكتب والدوريات ، وتجاهل ذلك عند الترجمة (مثلا: "يذكر بلوم في انغلاق العقل الأمريكي" بدلا من يذكر بلوم في كتابه "انغلاق العقل الأمريكي").

٢-١-٣-٣-تعليقات على استخدام الأخطاء في تقويم الترجمة:

عند النظر في الأخطاء اللغوية، يمكننا أن نصنفها من حيث أهميتها في ضوء عدة معايير، أهمها مايلي:

- أخطاء كلية وأخطاء جزئية (محلية local).
- أخطاء تؤثر في الفهم وأخطاء تسبب الإزعاج للقارىء.
- أخطاء تتعلق بشيوعها أو شيوع القواعد التي تخالفها، وأخطاء غير متكررة أو أخطاء لاتخالف قواعد شائعة في اللغة. (انظر مثلا: Carl James).

كذلك يمكننا أن نصنف الأخطاء من حيث نوعها اللغوي إلى عدة فئات، يهمنا منها مايلي:

أولا -أخطاء نحوية تركيبية:

أ- تؤثر في معنى الجملة والنص، مثل تغيير مكان الفاعل أو المفعول به مما يعكس الوظيفة النحوية للكلمة (خاصة في غياب التشكيل). كذلك مطابقة الصفة لكلمة غير الموصوف الحقيقي في الجملة (كما في عبارة "مدير المحطة الجديد." فهي صحيحة إن كان الموصوف هوالمدير وليست المحطة ولكنها خاطئة إن كان المقصود بالجديد المحطة.) ومن أمثلة ذلك أيضا ما يسمى الاستعمال المعكوس لبعض الأفعال (مثل "يعجبني التلميذ المجتهد" مقابل "أعجِب التلميذ المجتهد" و"فاتني القطار" – ينظر مقابلاتها في الإنجليزية لفهم المقصود من هذه الملاحظة.)

ب- قد لاتؤثر على المعنى على الرغم من إخلالها بقواعد اللغة، مثل غياب بعض جوانب المطابقة بين الصفة والموصوف (من حيث التذكير والتأنيث والعدد) مالم يكن هناك أكثر من موصوف واحد. ومثل ذلك غياب المطابقة بين الفعل والفاعل من حيث التذكير والتأنيث والعدد، مثل "العمال يحفر"، فهذه كلها لاتعيق فهم المعنى المقصود من الجملة (وهو ما نسمعه عادة في لغة العجم عندما يتحدثون بالعربية الهجين).

ثنيا: أخطاء معجمية:

أ- استخدام الألفاظ الخاطئة.

ب- استخدام التعبيرات الاصطلاحية الخاطئة أو ترجمتها حرفيا.

ت- الخطأ في المصاحبات اللفظية، مثل "رأب الكسر" بدلا من "رأب الصدع"

مما لاشك فيه أن الأخطاء المعجمية تعيق الفهم أو تؤدي إلى فهم خاطئ للمقصود، فهي التي تحمل العبء الدلالي الأساسي للجملة.

ثالثا: أخطاء تداولية:

يقصد بذلك استعمال الجملة أو العبارة في غير محلها، من حيث المقام، مثل عبارة "فضيلة مدير المدرسة " (مالم يكن شيخا فعلا). وهذا النوع قد لايعيق الفهم، ولكنه يسبب إزعاجا أو تسلية للقارئ أو السامع.

رابعا: أخطاء في ربط الجمل:

أ- عدم الربط بين الجمل والفقرات في النص العربي (مثل و، ف، ثم، ومن ثم...) نظرا لاكتفاء بعض اللغات مثل لإنجليزية بعلامة الوقف – النقطة- دون الحاجة إلى أداة ربط في الجملة التالية.

ب استعمال أداة الربط الخاظئة بين الجمل أو العبارات، مثل "بناء على ذلك، من ثم، لأن، لكن..." ولاشك أن مثل هذا الخطأ قد يؤدي إلى خطأ في الفهم أو إلى صعوبة فيه.

خامسا: أخطاء شكلية:

تشمل الأخطاء الإملائية وبعض الأخطاء في علامات الترقيم. ولكن يلاحظ أن بعض الأخطاء الإملائية قد تؤدي إلى خطأ معجمي.

٢ -٢- نظام BLEU الآلي لتقييم الترجمة الآلية:

يعتبر هذا البرنامج من أشهر االبرمجيات لتقييم الترجمة الآلية حاسوبيًا. ويتم استخدامه بإعداد ترجمة بشرية احترافية للنص وترجمة آلية له. ثم نخضع الترجمتين لنظام بلو الذين يقارن مستوى الدقة في الترجمة على مستوى الجمل لكل من الترجمة البشرية والآلية. ويقدم تقييمًا للترجمة الآلية في ضوء ذلك. (انظر Papineni في قائمة المراجع.)

٢-٣- مقترحات لتنقيح النصوص المترجمة آليا:

حيث إنه ثبت أن لاتوجد ترجمة عالية المستوى لنظام تام الآلية Fully Automatic High Quality Translation (FAHQT)، كما ذكر أحد منتقدي الترجمة الآلية منذ عدة عقود. وأنه باستثناء أنظمة الترجمة المتخصصة (مثل نظام ترجمة أخبار الطقس في كندا Meteo ونظام ترجمة الدراسات الخاصة بالنسيج Titus IV)، لابد للنص المترجم آليا أن يخضع للتنقيح والمراجعة، فنورد أدناه المعايير الخاصة بذلك التي أوردها برايان موسوب Brian Mossop في كتابه (التنقيح والمراجعة للمترجمين *Editing and Revising for Translators*)، هي:

١- الدقة (الأمانة في نقل المعنى المقصود)
٢- الاكتمال (عدم الزيادة أو النقصان)
٣- المنطق (غياب المتناقضات والمستحيلات العقلية)
٤- الحقائق (صدق الوقائع والمعلومات)

٥- السلاسة (في الأسلوب وسهولة قراءة النص)
٦- المناسبة للقراء والمستفيدين (تكييف النص المترجم لنوعية القراء الموجه لهم).
٧- اللغة الخاصة (مناسبة اللغة لنوع النص)
٨- الجانب العرفي (مراعاة السلامة الأسلوبية والاستخدام اللغوي)
٩- الآليات (الصحة النحوية والإملائية وفي علامات الترقيم...)
١٠- الشكل العام (توزيع الفقرات والجداول وتنظيمها)
١١- الطباعة واستخدام الأبناط
١٢- التنظيم (ترتيب الفقرات والهوامش والإحالات وترقيمها وطريقة كتابة العناوين الرئيسة والفرعية...).

مما لاشك فيه أن هذه المعايير تنطبق على كل من الترجمة البشرية والآلية، ولكن أهميتها أكبر في النصوص المترجمة آليا نظرا لكثرة الأخطاء الواردة فيها وللحاجة إلى تكييفها كي تلائم الغرض الذي تتم له الترجمة، مما لايمكن أن يقوم به سوى المترجم البشري المتمرس.

الخاتمة:

تحدثنا فيما سبق عن وسائل وأدوات حاسوبية تخدم الترجمة والمترجمين، بدءًا بوسيلتي الترجمة الآلية وذاكرة الترجمة، ثم معينات الترجمة مثل قواعد بيانات النصوص المترجمة والمعطيات المصطلحية والمعاجم الإلكترونية، ثم معينات المترجمين الآخرى مثل برامج معالجة النصوص والتدقيق الإملائي والنحوي والشابكة العنكبوتية والإملاء الآلي ووسائل التواصل عن بعد.

ويتضح لنا من العرض الموجز أعلاه أن للحاسوب دورا كبيرا في خدمة الترجمة والمترجمين، وأنه لايمكن لأي مترجم أو مؤسسة معنية بالترجمة أن تستغني عن هذه الوسائل والمعينات بشكل أو خر. بل إن هناك مقترحات لإيجاد ما يسمى بمحطة الترجمة translation station، حيث تجمع هذه الخدمات في مكان واحد لخدمة المترجم، يستعين بأيها شاء وفق احتياجاته، مما سيوفر عليه الكثير من الوقت والجهد.

أما بخصوص الترجمة الآلية فإننا لا نستطيع أن نعطي رأيا قاطعا في إمكاناتها، لأن ذلك يعتمد على عدة عوامل، منها (١) الغرض من الترجمة (٢) نظام الترجمة الآلية المستخدم و(٣) طبيعة عمل المترجم أو

المؤسسة التي يعمل بها و(٤) حجم العمل ونوعه، بل و(٥) اتجاه الترجمة. غيرأننا نستطيع القول بأن مشكلات الترجمة من العربية إلى اللغات الأخرى غالبا ماتكون أكبر منها في الترجمة من اللغات الأجنبية إلى اللغة العربية، ولعل من أهم أسباب ذلك غياب التشكيل في النصوص العربية، مما يجعل المرحلة الأولى من عملية الترجمة (فهم النص) بالنسبة للحاسوب عملية لا تخلو من التعقيد، فالكلمة الواحدة تقرأ بطرق متعددة، مما يؤثر سلبا في فهم الجملة قبل ترجمتها. كذلك من المعروف أن الترجمة الآلية في هذا الاتجاه حديثة نسبيا، إذا ماقارناها بالترجمة من اللغات الأجنبية (الإنجليزية خاصة) إلى اللغة العربية.

المراجع

الحميدان، عبدالله (١٩٩٨) الترجمة الآلية. في: **ندوة** تعميم التعريب...، ٣٢٣-٣٧٦.

رشوان، محسن والسعيد، المعتز بالله (٢٠١٩) **مقدمة في حوسبة اللغة العربية**. الرياض: مركزالملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

صيني، محمود إسماعيل (١٩٨٦) "الترجمة**:** معيناتها ووسائلها": **مجلة الفيصل** (الرياض)ع١٢١، ٤٠-٤٣.

صيني، محمود إسماعيل (١٩٨٩) "المعاجم في الترجمة الآلية" في**: اللسانيات العربية والإعلامية**. تونس:مركز الدراسات الاقتصادية والاجتماعية،، ١٩٦.-١٨٣

صيني، محمود إسماعيل (١٩٩١) "الترجمة الآلية: إمكاناتها وحدودها" في:**المجلة العربية للثقافة** (تونس)، ع ٢١ (سبتمبر) ١٩٩١م، ١٤٤.-١٣٢

صيني، محمود إسماعيل(١٩٩٤) "بنوك المصطلحات الآلية والمعاجم الألكترونية": **السجل العلمي للندوة الثانية لتعريب الحاسوب** (٣ -٢٧. مارس ١٩٩٤م)، ٣٣٤.-٣١١

صيني، محمود إسماعيل (١٩٩٤) "الترجمة الآلية واللغة العربية**:" مجلة التواصل اللساني** (فاس) مج ٦، ع١-٢، ١٩٩٤م، ٨٧.-٨٣

صيني، محمود إسماعيل (١٩٩٦ أ) "الحاسوب والترجمة": مجلة التواصل اللساني (فاس) مج٣، ١٩٩٦م، ٩١-،١.

صيني، محمود إسماعيل (١٩٩٦ب). "التطورات الحديثة في الترجمة الآلية" في: ندوة التعريب والحاسوب. دمشق: الجمعية العلمية السورية للمعلوماتية ، ١٩٩٦م، ١٦٩.-١٤٤

صيني، محمود إسماعيل (١٩٩٩ أ) "الحاسوب في خدمة الترجمة والتعريب ("(١/٢: **الثقافية** (لندن)، س ٦، ع ٢. (أبريل/ مايو ١٩٩٩م)، ٧١.-٦٨

صيني، محمود إسماعيل (١٩٩٩ ب) "الحاسوب في خدمة الترجمة والتعريب ("(٢/٢: **الثقافية** (لندن)، س ٦، ع ٢١ (يوليو/ أغسطس ١٩٩٩م)، ٧٦.-٧٢

صيني، محمود إسماعيل (١٩٩٩) "بنوك المصطلحات الآلية" في: **اللسان العربي**، العدد ٤٨، ١٩٩٩، ١-١٧.

صيني، محمود إسماعيل (١٩٨٦) وعبدالرزاق العاني (تحرير). **دراسات في الترجمة الآلية (أبحاث ندوة)**. الرياض: مدينة الملك عبد العزيز للعلوم والتقنية.

صيني، محمود إسماعيل **(٢٠٠٠)** "الاتجاهات المعاصرة في حركة الترجمة في العالم" في: مركز دراسات الوحدة العربية. **الترجمة في الوطن العربي:نحو إنشاء مؤسسة عربية للترجمة (بحوث ومناقشات الندوة الفكرية التي نظمها مركز دراسات الوحدة العربية).** بيروت: المركز،صص ١٧٢.-١٣٥

عبده، داود (٢٠١٠). **في اللغة والحاسوب: الترجمة وتدقيق الإملاء بين الإنسان والآلة**. عمان: دار جرير للنشر والإعلام.

عفيفي، إلهامي (١٩٩٨) "آفاق الترجمة الآلية وقصورها" في: **ندوة تعميم التعريب**...، ٣٧٧-٤٠٤.

ندوة تعميم التعريب وتطوير الترجمة في المملكة العربية السعودية المعقودة في رحاب جامعة الملك سعود في الفترة من ٢-٣ جمادى الآخرة، ١٤١٩هـ (٢٢-٢٣ ديسمبر ١٩٩٨م.) الرياض: جامعة الملك سعود.

المراجع الأجنبية:

James, Carl ( ). *Errors in Language Learning and Use: Exploring Error Analysis.* London and New York: Longman

Masterman, M (٢٠٠٣). "Mechanical Pidgin Translation" in: Nirenberg, S., Somers and Wilks, Y.(eds., ٢٠٠٣). *Readings in Machine Translation*. Cambridge, MA: MIT Press, ١٧٧-١٩٢.

Mossop, Brian (٢٠٠٧) *Editing and Revising for Translators. 2nd Edition.* Manchester: St Jerome Publishing.

Nagao, M, Tsujii, J and Nakamura, J (١٩٨٨). "The Japanese Government Project for Machine Translation" in: Slocum (Ed.) (١٩٨٨). *Machine Translation Systems*. Studies in Natural Language Process. Cambridge: CUP, ١٤١-١٨٦.

Papineni, Kishore, Salim Roukos, Todd Ward, and Wei-Jing Zhu “BLEU: a Method for Automatic Evaluation of Machine Translation”. *In Proceedings of the ٤٠th Annual Meeting of the Association for Computational Lin}guistics (ACL*), Philadelphia, July ٢٠٠٢, pp. ٣١١-٣١٨.

Sieny, Mahmoud I (١٩٨٦) "BASM: The Saudi Terminology Data Bank." In: Sieny and Abdul-Wahhab (eds, ١٩٨٦). *Studies in Machine Translation.* Riyadh: KACST.

Trujillo, Arturo (١٩٩٩). *Translation Engines: Techniques for Machine Translation*. London and Berlin: Springer-Verlag.

الذكاء الاصطناعي والترجمة من العربية وإليها

المقدمة:

يلاحظ المتابع للتطور العلمي في العقود الأخيرة انفجارًا كبيرًا في كمية المعلومات ونوعها في لغات مختلفة، مما جعل من الاستحالة بمكان ملاحقتها عن طريق أساليب الترجمة التقليدية (البشرية). ويشهد على ذلك أن قوقل تذكر أنها تترجم ١٠٠ مليار كلمة يوميًا!

فكان الحل في الاستفادة من الحاسوب بقدراته الهائلة في التعامل مع الكثير من المشكلات. فقد شهد النصف الثاني من القرن العشرين والعقدين الأولين من القرن الحادي والعشرين تطورا مذهلا في مجال الحواسيب والبرمجيات الحاسوبية كما وكيفا، وقدمت بذلك إمكانات وخدمات كبيرة للعاملين في شتى المهن وحقول المعرفة. ولا شك أن الترجمة لقيت الكثير من العناية في هذا المجال، بدءا بالمحاولات الأولية في العقد السادس من القرن الماضي لترجمة النصوص العلمية من الروسية إلى الإنجليزية.

وقد شاهدنا في الآونة الأخيرة تفجرا في أنواع المعينات التقنية، الحاسوبية خاصة، للمترجمين، على مستويات مختلفة، تبدأ بما يسمى بالترجمة الآلية (أو الترجمة بمعاونة الحاسوب) وانتهاء بمنسقات النصوص والإملاء الآلي automatic dictation وما بينهما من ذاكرات الترجمة وبنوك المصطلحات الآلية terminology data banks والمعاجم المحوسبة (أحادية اللغات وثنائيتها)، ثم الهواتف المحمولة.

سنتحدث فيما يلي عن الترجمة الآلية ودور الذكاء الاصطناعي في حل بعض مشكلاته.

١- مشكلات الترجمة الآلية من العربية المكتوبة وإليها مع نماذج حقيقية:

١-١- مشكلات الترجمة الآلية من العربية المكتوبة:

في الأمثلة التالية، نجد النص العربي متبوعًا بترجمته آليًا.

١-١-١- تعدد معاني بعض الألفاظ polysemy، وارتباط ذلك بالسياق (اللفظي الصغير micro-context والكبير macro-contextأو الموقف situation)أو الموضوع topic

بث القائد العيون بن صفوف العدو

The commander broadcast eyes in the ranks of the enemy

١-١-٢- المشتركات اللفظية homonyms وارتباط ذلك بالسياق (اللفظي الصغير micro-context والكبير macro-contextأو الموقف) أو الموضوع topic

لدي شعور قوي بأن الشعور المشوي في هذا المطعم أفضل من المقلي

I have a strong feeling that the roast feel in this restaurant is better than fried

١-١-٣- اتصال مورفيمات (الوحدات الصرفية) العربية مع بعضها البعض، مما يؤدي إلى وجود جمل كاملة في صورة كلمة واحدة (سألتمونيها = سألت+مو+ن+ي+ها =أنتم سألتموني إياها، فسيكفيكهم =ف+س+يكفي+ك+هم...)

God will suffice them فسيكفيكهم الله (God will protect you from them.)

١-١-٤- الاشتراك الكتابي homography (فهم= فـ+هُم، فـ+هَمَّ، فَهم، فَهِمَ، فُهِمَ)

فهم بها وهمت به لولا أن رأى برهان ربه

He understood it and dreamed of it had it not been for the proof of God’s proof

فهمت المسألة / فُهِمت المسألة.

I understood the issue

فهم الطالب الدرس

Understanding the lesson student

كان يعمل وراقا

He was doing paper

الوراقون معروفون في التاريخ العربي

Warraqis are well-known in Arab history

١-١-٥- غياب التشكيل، مما يفاقم مشكلة الاشتراك اللفظي و يؤدي إلى الاشتراك الكتابي وغموض إعراب الكلمات.

انتقَلَ إلى جوار الكعبة (انتقِل إلى جوار الكعبة)

Go next to the Kaaba

١-١-٦- الإعراب وأثره في الوظيفة النحوية.

دخل المدرسةَ طلاب كثيرون (مع نصب كلمة المدرسة)

Entered the school many students.

٢- الضمير العائد إلى اسم خارج الجملة (في جملة / جمل أخرى)، في وجود أكثر من اسم يعود إليه الضمير.

٣- الاستعمال المجازي للألفاظ والعبارات

المثال:

بث القائد العيون بن صفوف العدو

The commander broadcast eyes in the ranks of the enemy

١-١-٧- الأمثال والتعبيرات الاصطلاحية idioms

رجع بخفي حنينHe came back invisibly Hanin

ضرب أخماسًا في أسداس He hit a fifth in a sixth

انتقل إلى جوار ربه He moved next to his Lord

<في تجربة لبعض أنظمة الترجمة الأخرى، و جدنا: He came back with two daggers, Back in hiding لترجمة رجع بخفي حنين. وجاءت .Go to the god of Gods ترجمة لعبارة "انتقل إلى جوار ربه".>

١-١-٨- أسماء العلم، حيث إن كثيرًا منها لها معنى لغوي (فتح الباب، خضراء...)
الدكتور فتح الباب Dr. Open door السيدة سلمى خضراء Mrs. Salma Green

(لمزيد من الأمثلة، انظر عبده وعفيفي في الفصل السابق)..

هناك مشكلات أخرى كثيرة في الترجمة إلى العربية من الإنجليزية من أسبابها:

١-٢- التذكير والتأنيث للكلمات التي لاتفرق بينها بعض اللغات (مثل teacher التي قد تكون معلم أو معلمة).

١-٣- قواعد المطابقة:

- الصفة والموصوف: تعريف وتنكير، تذكير وتأنيث، العدد، الإعراب (ويلاحظ معاملة جمع غير العاقل مثل المفرد المؤنث – مدارس كثيرة)

الفتيات الذكية smart girls

- الاسم والضمير العائد إليه: التذكير والتأنيث، العدد (مع ملاحظة معاملة جمع غير العاقل مثل المفرد المؤنث، كما هو الأمر مع الصفة).
- الاسم وضمير الصلة: التذكير والتأنيث، العدد (مع ملاحظة معاملة جمع غير العاقل مثل المفرد المؤنث، كما هو الأمر مع الصفة). Women who النساء الذين

.These ladies have many children

هؤلاء السيدات لديهم العديد من الأطفال.

١-٤- هذا إلى جانب المشكلات العامة بين اللغات: مثل تعدد المعاني، الاشتراك اللفظي، التعبيرات الاصطلاحية، الاستعمال المجازي للألفاظ والعبارات، الإحالة reference)

مثال على الاشتراك اللفظي

The fans are new but the irons are old

المعجبون جدد، لكن المكواة قديمة

مثال على التعبير المجازي:

Power hungry.

المتعطشين للسلطة

They are power hungry.

إنهم قوة جائعة.

مثال على التعبير الاصطلاحي.

This does not ring a bell.

هذا لا يرن الجرس

(لمزيد من أمثلة مشكلات الترجمة الآليه من الإنجليزية إلى العربية، انظر القسم الخاص بتحليل الأخطاء في الفصل السابق: "التقنية في خدمة الترجمة".

٢- أساليب التعامل مع المشكلات المشار إليها أعلاه:

٢-١- التحرير المسبق للنصوص

٢-٢- الترجمة التحاورية، حيث يتم التعاون بين المترجم والآلة.

٢-٣- التحرير اللاحق.

ويلاحظ في هذه الأساليب دور الذكاء البشري، أي المترجم البشري الذي يتدخل إما قبل الترجمة الآلية أو أثناءها أو بعدها. وقد أثبتت هذه الأساليب تكلفتها المالية وتوافر المترجم البشري المطلوب ومن حيث الوقت المطلوب للترجمة (وهو ما يعتبرحاسمًا في ظروف الحاجة إلى الكميات الهائلة من النصوص المترجمة في ظروف الانفجار المعلوماتي).

٢-٤- الأسلوب الإحصائي: وهذا الأسلوب يتطلب وجود مدونات لغوية متوازية parallel corpora كبيرة جدًا. ومع ذلك فلا تزال محدودة بشكل أوبآخر، من حيث الجودة والتنوع والأسلوب. من هنا

كان التفكير في الأسلوب الخامس الترجمة الآلية العصبية neural machine translation، واللجوء إلى الذكاء الاصطناعي.

٢-٥- الذكاء الاصطناعي:

يلخص بعضهم مفهوم الذكاء الاصطناعي بأنه نظام يحاول أن يحاكي الذكاء البشري في معالجة المشكلات التي تواجهه مستفيدًا من القدرة على الاستنتاج والتعلم المستمر.

٣- دور الذكاء الاصطناعي

٣-١- أنظمة الترجمة الحديثة:

إن نظرة سريعة على أنظمة الترجمة الحديثة تنبؤنا أن هناك أربعة أنظمة معروفة هي:

٣-١-١- نظام الترجمة الألية المبني على الإحصاء Statistical Based MT: النظام الذي يعتمد على نظرية الاحتمالات في ضوء المعطيات المتوافرة، خاصة من المدونات اللغوية المتوازية.

٣-١-٢- نظام الترجمة الآلية المبنية على القواعد Rule Based MT وهو النظام الذي يعتمد على تقعيد الظواهر اللغوية. ونحن نعرف أن ليس كل ظاهرة لغوية يمكن تقعيدها. فحتى في كتب النحو المعروفة نجد ما يسمى بالسماع، أي الاستثناء من القاعدة. فكيف بالألفاظ وتنوعاتها ومعانيها واستعمالاتها المجازية، إضافة إلى ما أشرنا إليه أعلاه من تعبيرات اصطلاحية وما شابهها؟

٣-١-٣- نظام الترجمة الآلية الهجينة Hybrid MT من النوعين السابقين: لا شك أن هذا اعتراف بعدم كفاية أي من الأسلوبين المذكورين، حيث وجد الباحثون ضرورة الاستفادة منهما كليهما.

٣-١-٤- نظام الترجمة العصبوني Neural MT بنوعيه التقليدي والمطور. (انظر: Localize – Blog (٢٢ June, ٢٠١٨.

يعرف بعضهم النظام العصبوني بأنه: "ينتمي إلى عائلة تشفير-فك التشفير enode-decode للجملة إلى متجه vector ذي طول ثابت." ويتصف بقدرته على "الفهم العميق" والتعلم من الأمثلة التي ترد إليه. ويرى بعض الباحثين أن هذا صورة مطورة للترجمة الآلية المبنية على الإحصاء.

وفي أحد التصاميم المجددة يذكر الباحثون بأن التجديد يكمن في نظام ثنائي الاتجاه وفي مرونة طول المتجه الثابت سابقًا. انظر على سبيل المثال:

Bahdanau, D. and Cho, K. (٢٠١٥) ; Bentivogli, L, Bisazza, A, Cettolo, M, Federicao, M. (٢٠١٦).

وهذا النظام الأخير يبدو أنه يحاول الاستفادة من إمكانات الذكاء الاصطناعي في حل مشكلات الترجمة التي ألمحنا إلى بعض منها فيما سبق. فمن خلال الذكاء الاصطناعي قد نجد حلولًا للمشكلة الكبرى في فهم النصوص، وهو الخطوة الأولى والأساس لأية ترجمة. فربما استطعنا أن نعالج مشكلة فهم الكلمة أو التعبير في السياقات المختلفة، سواءً أكانت لفظية أم متعلقة بالموقف.

من هنا أيضًا قد نجد في الذكاء الاصطناعي حلولًا لمشكلات تعدد معاني الألفاظ والتعابير، والمشتركات اللفظية، والمشتركات الكتابية (خاصة في العربية)، إضافة إلى مشكلات التعرف على معاني التعابير الاصطلاحية (بما في ذلك الأمثال) والتمييز بين الاستعمال الاصطلاحي وغير الاصطلاحي لكثير من التعابير، مما أوردنا في الأمثلة على المشكلات التي تواجه الترجمة من العربية أو إليها. (انظر Tan, Z et al. (٢٠٢٠).

كلمة أخيرة:

في الصفحات السابقة حاولنا إعطاء تصور مبسّط لعمليات الترجمة الألية وأنواعها وذكرنا أمثلة لبعض أنظمة الترجمة الآلية التي تتعامل مع اللغة العربية.

ثم عرجنا على أهم المشكلات اللغوية التي تواجه الترجمة الآلية من العربية وإليها، مع التمثيل لها من واقع ترجمات جربناها على أحد الأنظمة المتطورة في الترجمة، نظام قوقل للترجمة Google Translate.

ونقول في خاتمة حديثنا إن كل من خبر الترجمة أو مارسها يدرك أن الترجمة تجمع بين العلم والفن وبين القواعد والسليقة اللغوية. لذلك لا نستطيع أن نحدد بدقة ما يقوم به أي مترجم محترف، من ثم نعلمه للحاسوب ليحاكيه. فكل "شيخ له طريقة" كما يقال. ولكن هناك مشكلات لغوية وتقنية يمكننا أن نتعامل معها سواءً من خلال القواعد أو الاحتمالات أو تعلُّم الآلة.

هنا يرد السؤال الملح علينا بوصفنا مترجمين أو مستفيدين من الترجمة: هل ستحل الآلة يومًا مكان المترجم الجيد أو المحترف؟ الجواب: كلا مالم يصبح الذكاء الاصطناعي ذكاءً بشريًا. والمستقبل المنظور ينبؤنا بأن الآلة لن تحل مكان المترجم ما لم يكن مترجمًا آليا، لا يفكر ولا يعي ما يترجم، كما نلاحظ في كثير من الأحيان.

المراجع

جادالكريم، عبدالله (١٤٤١هـ/ ٢٠٢٠م) **عبقرية اللغة العربية والذكاء الاصطناعي: آفاق وطموحات.**

طنطا (مصر): دار النابغة للنشر والتوزيع.

الحميدان، عبدالله (١٩٩٨) الترجمة الآلية. في: **ندوة** تعميم التعريب...، ٣٢٣-٣٧٦.

السجل العلمي للندوة الثانية لتعريب الحاسوب (٣ -٢٧. مارس ١٩٩٤م)

صالح، محمود إسماعيل (٢٠١٨) التقنية في خدمة الترجمة والمترجمين، في كتاب: **اللسانيات الحاسوبية واللغة العربية: إشكالات وحلول** (تحرير د/ عمر مهديوي)، عمان: دار كنوز المعرفة للنشر والتوزيع، صص ٥٣-٨٦.

صيني، محمود إسماعيل (١٩٨٩) "المعاجم في الترجمة الآلية" في**: اللسانيات العربية والإعلامية**. تونس:مركز الدراسات الاقتصادية والاجتماعية، ١٩٨٩م، ١٩٦.-١٨٣

صيني، محمود إسماعيل (١٩٩١) "الترجمة الآلية: إمكاناتها وحدودها " في: **المجلة العربية للثقافة** (تونس)، ع ٢١ (سبتمبر) ١٩٩١م، ١٤٤.-١٣٢

صيني، محمود إسماعيل (١٩٩٤)" بنوك المصطلحات الآلية والمعاجم الألكترونية": **السجل العلمي للندوة الثانية لتعريب الحاسوب** (٣ -٢٧. مارس ١٩٩٤م)، ٣٣٤.-٣١١

صيني، محمود إسماعيل (١٩٩٤). "الترجمة الآلية واللغة العربية:" **مجلة التواصل اللساني** (فاس) مج ٦، ع١-٢، ١٩٩٤م، ٨٧.-٨٣

صيني، محمود إسماعيل (١٩٩٦). "الحاسوب والترجمة": **مجلة التواصل اللساني** (فاس) مج٣، ١٩٩٦م، ٩١-،١.

صيني، محمود إسماعيل (١٩٩٦). "التطورات الحديثة في الترجمة الآلية" في: **ندوة التعريب والحاسوب**. دمشق: الجمعية العلمية السورية للمعلوماتية، ١٩٩٦م، ١٦٩.-١٤٤

صيني، محمود إسماعيل.(١٩٩٩أ) "الحاسوب في خدمة الترجمة والتعريب ("**(**١/٢: **الثقافية** (لندن)، س ٦، ع ٢. (أبريل/ مايو ١٩٩٩م)، ٧١.-٦٨

صيني، محمود إسماعيل (١٩٩٩ب)**.** "الحاسوب في خدمة الترجمة والتعريب ("**(**٢/٢: **الثقافية** (لندن)، س ٦، ع ٢١ (يوليو/ أغسطس ١٩٩٩م)، ٧٦.-٧٢

صيني، محمود إسماعيل وعبدالرزاق العاني (تحرير، ١٩٨٦). **دراسات في الترجمة الآلية (أبحاث ندوة)**. الرياض: مدينة الملك عبد العزيز للعلوم والتقنية.

صيني، محمود إسماعيل **(٢٠٠٠)** "الاتجاهات المعاصرة في حركة الترجمة في العالم" في: مركز دراسات الوحدة العربية. **الترجمة في الوطن العربي:نحو إنشاء مؤسسة عربية للترجمة (بحوث ومناقشات الندوة الفكرية التي نظمها مركز دراسات الوحدة العربية).** بيروت: المركز،صص ١٧٢.-١٣٥

عبده، داود (٢٠١٠). **في اللغة والحاسوب: الترجمة وتدقيق الإملاء بين الإنسان والآلة**. عمان: دار جرير للنشر والإعلام.

عفيفي،، إلهامي (١٩٩٨) "آفاق الترجمة الآلية وقصورها" في: **ندوة تعميم التعريب**...، ٣٧٧-٤٠٤.

ندوة تعميم التعريب وتطوير الترجمة في المملكة العربية السعودية المعقودة في رحاب جامعة الملك سعود في الفترة من ٢-٣ جمادى الآخرة، ١٤١٩هـ (٢٢-٢٣ ديسمبر ١٩٩٨م.) الرياض: جامعة الملك سعود.

المراجع الأجنبية:

Bahdanau, D. and Cho, K. (٢٠١٥) "Neural Machine Translation by Jointly Learning to Align and Translate". At: https://arxiv.org/abs/١٤٠٩,٠٤٧٣

Bentivogli, L, Bisazza, A, Cettolo, M, Federicao, M. (٢٠١٦) "Neural versus Phrase-Based Translation Quality: A Case Study." **www.aclweb.org/anthology/D١٦-١٠٢٥.pdf**

Cho, K., van Merrienboer, Bahdanau, D, and Bengio, Y. (٢٠١٤) “On the Properties of Neural Machine Translation: at: https://www.researchgate.net › publication › ٢٦٥٣٨٥٨٧٩_...

Localize – Blog (٢٢ June, ٢٠١٨) “Different Types of Machine Translation”

Masterman, M (٢٠٠٣). "Mechanical Pidgin Translation" in: Nirenberg, S., Somers and Wilks, Y.(eds., ٢٠٠٣). *Readings in Machine Translation*. Cambridge, MA: MIT Press, ١٧٧-١٩٢.

Mossop, Brian (٢٠٠٧) *Editing and Revising for Translators. 2nd Edition*. Manchester: St Jerome Publishing.

Nagao, M, Tsujii, J and Nakamura, J (١٩٨٨). "The Japanese Government Project for Machine Translation" in: Slocum (Ed.) (١٩٨٨). *Machine Translation Systems. Studies in Natural Language Process*. Cambridge: CUP, ١٤١-١٨٦.

Sieny, Mahmoud I (١٩٨٦). "BASM: The Saudi Terminology Data Bank" in: Sieny, Mahmoud and Ani (eds) *Studies on Machine Translation* (Conference Proceedings). Riyadh: King Abdulaziz City for Science and Technology, ١٩٨-٢١٩.

Tan, Z. et al. (٢٠٢٠). Neural machine translation: A review of methods, resources, and tools. http://www.keaipublishing.com/en/journals/ai-open

Trujillo, Arturo (١٩٩٩). *Translation Engines: Techniques for Machine Translation*. London and Berlin: Springer-Verlag.

Chung, J, Gulcehre, C., Cho, K. and Bengio, Y. (٢٠١٤)
Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. arXiv.١٤٠٩,١٢٥٩v٢

التعريب الآلي للأسماء الأجنبية: الأسس اللغوية

من الأمور التي تهم كل دولة ومؤسسة عربية أن يكون لها نظام يعينها في كتابة الأسماء الأجنبية بحرف عربي بصورة سليمة علميا ولغويا ومطردة عمليا، بحيث لا تخضع العملية للاجتهادات الشخصية بخلفيات المعرب اللغوية والعلمية واللغوية المختلفة، فتأتي الأسماء المعربة بصورة مضطربة بأشكال وصور عديدة يخيل إلى قارئها تعدد الأسماء المنقولة، وبطريقة يصعب على صاحب الاسم نفسه أن يتعرف عليه لبعده عن الأصل.

١- القضايا والأسس العلمية اللغوية لتعريب الأسماء:

١-١- الكتابة الصوتية أم النقحرة:

من الأمور التي قد تبدو بدهية للعاملين في مجال الترجمة أننا عندما ننقل الأسماء من لغة إلى أخرى كل ما علينا هو النقحرة: أي كتاب حرف عربي مقابل الحرف الأجنبي. غير أن الأمر ليس بهذه البساطة حتى بالنسبة للغات التي تستخدم نظام الكتابة الألفبائي كالإنجليزية والروسية والتركية. فالحرف اللاتني الواحد مثلا ينطق بصورة مختلفة في اللغة الواحدة أحيانا كما في الإنجليزية مثلا، كما أنه قد يكون صامتا كما هو الحال في حرف e في أواخر الكلمات الإنجليزية والفرنسية والصوامت كذلك في أواخر الكلمات الفرنسية. وسنتحدث عن هذه القضية مفصلا أدناه.

من هنا نجد أن المترجمين في واقع الأمر يقومون بالكتابة الصوتية بالحرف العربي للأسماء الأجنبية، مع اختلاف في نقل الأسماء تبعا لخلفية المترجم العربي من جهة: لبناني/سوري، عراقي، مصري، سعودي، مغاربي...إلخ (كما في نقل صوت g إلى غ و ك وج و ربما ق). وتبعا لخبرة المترجم وأسلوبه في التعامل

مع الصوائت vowels القصيرة، كالفتحة أو الضمة أو الكسرة التي لا تكتب عادة في الكتابة العربية المعتادة. فنجد أن البعض قد يعوض الفتحة بالألف والضمة بالواو والكسرة بالياء، مثل بوش وكلينتون مقابل Bush, Clinton اللذان ينطقان أصلا بضمة قصيرة وكسرة قصيرة.

١-٢- ما هي أسس النقل؟

أ- في ضوء ماذكرنا أعلاه، يبدو أن المتبع هو الكتابة الصوتية للاسم بالحرف العربي وليس مجرد النقحرة transliteration بمعناها الدقيق الذي ذكرناه أعلاه.

ب- من القواعد المعروفة في الترجمة مدى إمكانية الترجمة العكسية. وهذا يعنى، في حالة تعريب الأسماء الأجنبية، مدى إمكانية إعادة كتابة الاسم بصورته الأجنبية الأصلية بحد أدنى من الخطأ.

ت- الاطراد consistency في التعامل مع الأسماء، فلا يكتب مثل اسم ما بشكل أحيانا وبشكل آخر أحيانا أخرى. ويمكننا تحقيق ذلك إذا أوجدنا قواعد مطردة لكتابة الأصوات الأجنبية بالحرف العربي مع التفريق ما أمكن بين الأصوات والحروف المختلفة. مثلا لا يمكن للكتابة العربية التقليدية أن تفرق بين الزوجين التاليين: p x b, f x v فلا مانع من كتابة الأسماء التي تشتمل عليها بالباء والفاء. وكذا الفرق بين صوت الجيم المعروفة والجيم المعطشة (كما في نطق j في الإنجليزية والفرنسية) يمكننا التغاضي عنه عند التعريب. أما الاختلاف بين الجيم والجيم القاهرية فلا بد من مراعاته لأننا لدينا أساليب مقترحة لذلك (كتابة الأولى جيما والثانية غينا مثلا) فنفرق في التعريب بين Jordon x Gordon، خاصة وأن جميع اللغات الأوربية تقريبا تميز بين الصوتين ولها قواعد كتابية خاصة للتمييز بينهما، والتي تتمثل باختصار في نطق حرف g جيما إذا كان متبوعا بالحروف I, e, y ونطقه جيما قاهرية في الحالات الأخرى، بالإضافة إلى كتابة الجيم في صورة J في بعض اللغات الأوربية.

ث- لأسباب عملية (كالتعامل الحاسوبي – لوحة المفاتيح العادية وبرامج قواعد البيانات - والقراءة من غير المتخصصين) لا يسمح باستخدام رموز أو حروف غير الحروف العربية التقليدية، أي لا نستخدم مثلا رمز الجيم المعطشة (ج بثلاثة نقط چ) أو الفاء المجهورة (فاء بثلاثة نقط ڤ) أو الكاف الفارسية (الكاف التي عليها شرطة گ) أو الباء المهموسة (ذات النقط الثلاث پ).

ج- لأسباب علمية وعملية تعطى الأفضلية للتعريب الشفاف (الذي يسهل إرجاعه إلى أصله الأجنبي) على التعريب غير الشفاف (الذي يصعب إرجاعه إلى أصله الأجنبي). ومن أمثلة ذلك تعريب الاسم Chavez إلى "شابيث"، وهو صحيح وفقا لأ ساليب نطق الحرفين v, z في سياقات إملائية معينة في إحدى اللهجات

الأسبانية الأوربية. ولكن التعريب إلى "شافيز" أعطي الأولوية لقربه من شكل الحروف الأصلية وكذلك شيوع اللهجة التي تنطقها بهذه الصورة (باقي أسبانيا وأمريكا الجنوبية).

٢ـ مشكلات التعريب للأسماء المكتوبة بالحرف اللاتيني:

لابد لنا من التنويه هنا إلى أن هناك فئتان من اللغات تستخدم فيها الحروف اللاتينية في جوازات سفرها، إحداها بصورة أصلية والأخرى بصورة مرومنة عن نظام كتابة آخر، كما هو موضح في الفئة الثانية

الفئة الأولى:

أـ لغات أوربا الغربية
ب- اللغات الأسيوية (التركية، الإندونيسية)
جـ- اللغات الأفريقية (الصومالية، السواحيلية، الأفريكان)

الفئة الثانية:

أـ الكتابة الغربية (السيريلية، مثل الروسية واليونانية)
ب- الكتابات الشرقية الأبجدية (مثل البنغالية والتاميل)
جـ- الكتابات الشرقية الكلمية والمقطعية (مثل الصينية واليابانية)

وهناك مشكلات خاصة تتعلق بنظام الرومنة المتبع للغات الفئة الثانية أعلاه، خاصة فيما يتعلق بالأصوات التي لاتجد تمثيلا لها في الحرف اللاتيني، مثل الخاء في اليونانية. وتكمن المشكلات في جانبين: أحدهما أن بعض هذه اللغات لها أكثر من نظام للرومنة، حتى من الناحية العلمية. وثانيهما أن بعض هذه اللغات كاللغة الصينية لاتستخدم النظام الألفبائي، من ثم لا تتكون الأسماء من حروف بل من مقاطع (حرف ساكن +حركة) أوكلمات كاملة. فالاسم "عبد الكريم" مثلا يصبح بعد الرومنة شيئا مثل Abudukarimu وهومما يسبب

إرباكا للمترجم غير المسلم فيعرب الاسم إلى أبودوكريمو. وهذا بالإضافة إلى استخدام بعض الحروف اللاتينية لرومنة أصوات تخالف المتبع عادة في اللغات الأخرى، مثل استخدام Q لصوت قريب من "تش" وكذا الحرفان zh، واستخدام x ليقابل صوت الشين و h مقابل الصوت خ (في نظام الرومنة المعروفة بالبن ين).

هناك مشكلات عديدة تنتج من الاختلافات التي تستخدم بها كثير من الحروف اللاتنية سواء أكان ذلك في اللغات الأوربية أم اللغات الأخرى من المجموعات الثلاث أعلاه. وقد ناقش الباحث بعضا من هذه الأمثلة في بحث قدم إلى ندوة خاصة بهذا الموضوع عقدته جامعة نايف للدراسات الأمنية منذ أربع سنوات تقريبا.

وأورد أدناه فقرات من الدراسة المذكورة (انظر الملحق):

٣- أهم مشكلات كتابة الأسماء الواردة في لغات أوربا الغربية:

من قراءة متأنية لجداول الحروف الساكنة والمتحركة في اللغات الأوربية (انظرالملحق)، يمكننا ملاحظة مايلي:

أ- على الرغم من أن جميع هذه اللغات تستعمل الحرف اللاتيني إلا أن هناك اختلافات في نطق بعض الحروف سواء أوردت منفردة (مثل j) أم مجتمعة مع حروف أخرى (مثل: th, st, gn, ci, ti) وغيرها. والمشكلة أكبر مع الصوائت (الحركات). فكما نلاحظ من جدول الصوائت الأساسية (الجدول٥ في الملحق)، أنه حتى هذه لم تسلم من الاختلافات في نطقها، وليس من صائت واحد له نطق واحد في جميع لغات أوربا الغربية، بل وفي اللغة الواحدة أحيانا كما في الإنجليزية.

ب- كما ذكرنا أعلاه، يختلف نطق هذه الحروف ليس فقط تبعا للغة التي تستخدم فيها، بل وتتأثر كذلك بعوامل أخرى مثل الحروف التي قبلها والتي بعدها، كما نرى من جداول ثنائيات أو مجموعات الحروف الساكنة والمتحركة (الجداول ٢و٣ و٤ و٦ و٧). كما إنها تخضع لقواعد إملائية وصوتية كذلك، مثل منع الإنجليزية ورود صوت انفجاري (مثل k, p) قبل ساكن آخر في بداية الكلمات، من ثم يصبح هذا الصوت "ساكتا" – لاينطق به، كما في know (نُو)، بينما تجيز لغات أخرى كالألمانية مثل هذه الظاهرة،

ومثل عدم نطق الإنجليزية لحرف e عندما ترد بعد حرف ساكن في آخر الكلمة وعدم نطق الفرنسية للحرف الساكن عندما يرد في آخر الكلمة أحيانا، إلى غير ذلك من الأمثلة التي لاحصر لها.

فلا غرو أنه حتى أهل اللغة يعجزون في كثير من الأحيان عن كتابة الأسماء غير المألوفة بصورة صحيحة. من ثم كثيرا مانسمع في البلدان الناطقة بالإنجليزية السؤال المعروف: "How do you spell it? How do you spell your name?" أي "كيف تتهجاه؟ أو كيف تتهجى اسمك؟"

جـ- هناك مشكلات لغوية تتعلق بنظام الكتابة العربية، من أهمها (١) عدم وجود مقابلات عربية لبعض الحروف اللاتينية، مثل p, v, g على الرغم من وجود مقابلات لها، مثل الباء والفاء المثلثتين (بثلاث نقط) والكاف الفارسية، مما تستعمله بعض اللغات الإسلامية، ومثل الحركات المائلة (الكسرة والضمة المائلتين، كما في Jane, John مثلا) والتي لارموز لها في العربية أو اللغات التي تكتب بالحرف العربي (حسب علمنا)، مما يؤدي إلى اختلافات أحيانا في نقل بعضها إلى العربية، كما في كتابة الاسم Jane (هل يكتب جاين أو جين؟).

(٢) عدم استعمال الحركات القصيرة (الفتحة والضمة والكسرة) في الكتابة العربية المعتادة، على الرغم من وجودها في اللغة العربية، مما يلجيء كثيرًا من معربي الأسماء الأجنبية إلى إضافة ألف أو واو أو ياء للتعويض عنها عند كتاب الأسماء الأجنبية أحيانا، مثل كتابة الاسم Bush في صورة (بوش، وليس بُش كما تُنطق) و Munich على شكل (ميونيخ، بدلا من ميونِخ) وهكذا.

(٣) اختلاف العرب في كتابة بعض الحروف اللاتينية التي لها مقابل عربي، مثل حرف t التي تكتب تاء وطاء أحيانا (طوني وطاكسي بدلا من توني وتاكسي)، علما بأن بعض العرب الأوائل فعلوا ذلك، كما في أرسطوطاليس وسقراط.

كذلك نجد أن صوت الكاف (التي تكتب أحيانا c بالحرف اللاتيني) تم تعريبها إلى ق، كما في موسيقى و سقراط (مقابلاتهما في الإنجليزية music, Socrates).

والأمر أسوء في حالة الحروف التي لايوجد لها مقابل في العربية، مثل g، كما في Gulliver فالشاميون يكتبونها /غ/ والمصريون يكتبونها /ج/ والعراقيون يكتبونها /ك/ وبعضهم في دول الجزيرة العربية يكتبها /ق/، كما في: اللغة الإنغليزية، الإنجليزية، الإنكليزية، الإنقليزية.

(انظر الملحق (١) "أساليب استخدام الحرف اللاتيني لمناقشة مفصلة للموضوع).

٤ـ مشكلات تعريب الأسماء الأجنبية:

تتمثل أهم المشكلات في تعريب الأسماء الأجنبية ما يلي:

١. أسماء العلم الجغرافية: لابد لنا أن ندرك أن كثيرا من الأسماء الجغرافية مقابلات عربية، من ثم لا بد من إعطاء المقابل العربي لها، مثل England, Germany, Norway التي لها مقابلات متفق عليها: انجلترا، ألمانيا، النرويج). وينطبق ذلك بالطبع على أسماء المدن والبلدان والملامح الجغرافية العربية، مثل: Cairo, Damascus, Hebron وغيرها. وينطبق الأمر نفسه على مسميات المدن التاريخية الأسبانية مثل Seville, Toledo, Granada, Malaga حيث إن لها مقابلات عربية معروفة، هي إشبيلية وطليطلة وغرناطة والملقى. ويختلف الأمر عندما لايكون لهذه الأسماء خلفية تاريخية، مثل مدينية Toledo في ولاية أوهايو الأمريكية التي يجب أن تكون توليدو.
٢. نطق الحروف اللاتينية بصور مختلفة حتى في اللغة الواحدة أحيانا، وهو ما أشرنا إليه في مشكلات تعريب الأسماء المكتوبة بالحرف اللاتيني من الفئتين الأولى والثانية أعلاه. الأسماء الأجنبية الواردة في اللغات المختلفة (مثل الاسم الفرنسي أو الألماني الذي يرد ضمن الأسماء الأمريكية، وكذلك الأسماء الأوربية في الفلبين) والتي تحتفظ بشكلها ونطقها الأصلي ولا تخضع لقواعد الكتابة باللغة المضيفة.
٣. يدخل ضمن المشكلة الثالثة أعلاه بصورة واضحة الأسماء ذات الأصول الإسلامية أو العربية في جميع بلدان العالم المختلفة، حيث تتأثر كتابتها بأسلوب نطقها في تلك اللغات، بل قد تحتفظ بطريقة نطقها بلغة وسيطة أخرى (مثل كتابة اسم شخص مغربي الأصل بالطريقة الفرنسية، وهويحمل الجنسية الهولندية). فمثل هذه الحالات لايمكن تطبيق قواعد النقل من اللغة الأجنبية مباشرة، بل لابد من التعرف على النطق الأصلي والصحيح لها بالعربية.
ومن الأمثلة الطريفة نطق البنغالي الجيم العربية زايا، فنجد أن الاسم عبد الجبار مكتوبا Abdulzabbar وكذلك نطق الضاد والذال والظاء في دول شبه القارة الهندية زايا، كما في نطق الأسماء ضياء الحق، ذو الفقار وظفرالله.
٤. يلاحظ أن الرموز الإضافية التي تكتب فوق الحروف اللاتينية أو تحتها، مثل العلامات التي تكتب تحت أوفوق c, s في التركية والفرنسية لاتظهر في الصورة المحوسبة. ويمثل هذا مشكلة إضافة للمعرب، حيث

لابد أن يدرك أن c تحتها رمز معين ينطق سينا بالفرنسية وبرمز آخر في التركية يجعل الحرف جيما، وحرف s في التركية توضع له علامة إضافية لتجعله شينا.

٥. من المشكلات العملية المهمة خلفية المترجم (المعرب)، مثل اللهجة التي يعرفها (مثلا الأسبانية الجنوبية) أوخلفية المترجم (المعرب) للأسماء من اليونانية. ففي الحالة الأولى نجد أنه يعرب الاسم المعروف Chavez إلى "شابيث" بدلا من شافيز(وهو سليم لغويا في ضوء بعض لهجات اللغة الأسبانية). وفي المثال الثاني نجده يعرب الاسم Achmad إلى أشمد، بدلا من أحمد.

٥- الحلول المقترحة والمتبعة للمشكلات المذكورة:

أولا: من حسن حظ القائمين على المشروع أن مهمته تقتصر على تعريب أسماء العلم للأشخاص، ولا علاقة له بالأسماء الجغرافية أو أسماء العلم التاريخية.

ثانيا: كان هناك اتجاهان لتعريب الأسماء الأجنبية:

أ- الأسلوب الإحصائي ونظرية الاحتمالات.
ب- الأسلوب اللغوي الجغرافي.

٥-١- الأسلوب الإحصائي:

يرى أصحاب هذا الاتجاه أن نعتمد على مدونات لأسماء أجنبية مع مقابلاتها العربية، ويستخلص منها قواعد أو مقابلات وفق لنظرية الاحتمالات. وهو مايماثل توجه شركة آي بي إم في الترجمة الآلية بين اللغات. ولكن هذا الأسلوب له حدوده ومشكلاته التي نورد أهمها أدناه.

أولا: من المعلوم أنه يستحيل تطوير نظام شمولي لكل اللغات، نظرا للاختلافات الكثيرة في أساليب الرومنة فيها. وجدير بالذكر أننا قمنا بتجربة عملية لنظام من هذا القبيل (حسب الظاهر لنا) فثبت فشله بنسبة ٥٠% من الحالات.

ثانيا: يتطلب أي نظام إحصائي وجود عينة أو مدونة corpus يحتوي على عدد كبير جدا (مئات الآلاف) من الأسماء الأجنبية ومقابلاتها الصحيحة، وأود التأكيد على شرط "الصحيحة"، حتى يتسنى للنظام

الإحصائي أن يستخلص منها قواعد مناسبة. وهذا الشرط يصعب توافره إلا ربما فيما يتعلق بالأسماء لأشخاص من الولايات المتحدة والدول الناطقة بالإنجليزية. ولا شك أن أسلوب الاحتمالات الإحصائي هو الأنسب لمثل هذه الأسماء، نظرا لتنوع مصادرها ولنظام التهجئة الإنجليزية السيء الصيت، بشرط توافر الشرط المذكور أعلاه.

ثالثا: يلاحظ أن معظم اللغات لا تتوافر لدى الجهة القائمة بمشروع التعريب أعداد كافية من الأسماء مع مقابلات صحيحة لها. من ثم لايمكن تطبيق النظام الإحصائي عليها.

رابعا: هناك بلدان قد لاتتوافر حاليا لدى الجهة المذكورة أية أسماء لها، بله أن يكون لها مقابلات معربة.

٥-٢- الأسلوب اللغوي الجغرافي:

في ضوء المشكلات المذكورة أعلاه، وجدنا أن الأسلوب الأمثل هو تصميم برنامج من أنظمة فرعية تأخذ بعين الاعتبار المشكلات الخاصة باللغات المختلفة وأساليب الرومنة فيها، إضافة إلى مشكلة الأسماء العربية والإسلامية الأصل في هذه اللغات. فكان لابد من اتخاذ الخطوات العملية التالية:

١- حصر البلدان المختلفة (وفقا لعضوية الأمم المتحدة) والتعرف على لغاتها القومية.
٢- نظرا لتعدد اللغات في كثير من البلدان الأفريقية خاصة وبعض الدول الآسيوية (كالهند مثلا)، كان لابد من الاتفاق على عدد من اللغات الرئيسة لتكون منطلقا لدراسة أنظمة الرومنة فيها للوصول إلى نظام للتعريب المناسب. وكان الأسلوب المقترح هو جمع هذه البلدان في فئات، تبعا للغات القومية فيها واللغة المشتركة أو الرسمية في التعاملات الخارجية لهذه البلدان.
٣- من هنا تم اتخاذ القرار بتحديد اللغة الأوربية الرسمية أو شبه الرسمية في البلدان متعددة اللغات المحلية. من ثم اتفق أن تكون لغات المستعمر الأوربي السابق هي اللغة التي ينطلق نظام التعريب منها. فالدول العضو في الكومنولث اعتبرت الإنجليزية لغة لها والدول الفرانكوفونية كانت الفرنسية لغة لها وهكذا. والدول التابعة للاتحاد السوفيتي اعتبرت الروسية منطلقا منها للتعريب، مع استثناءات قليلة ربما أمكن اعتبار نظام الرومنة للتركية أساسا لتعريب أسماء مواطنيها.
٤- فيما يخص البلدان الآسيوية والأوربية الأخرى تم معالجة لغات كل منها بنظام مناسب لها، وذلك كالصينية واليابانية والإندونيسية والفلبينية وغيرها.

(انظر الملحق (٢) "مقترحات بخصوص البلدان التي وردت أسماؤها في القائمة الأخيرة")

٥- تم تطوير نظام خاص لكل لغة تم الاتفاق عليها، مراعيا للظروف اللغوية لها. وقد لجأت الشركة المنفذة لمشروع التعريب إلى مترجمين لهذه اللغات ليقوموا بتعريب الأسماء الواردة فيها، من ثم بناء نظام للتعريب لها.

٦- نظرا لأنه ثبت من التجربة أن الأسماء ذات الأصول العربية والإسلامية تمثل مشكلة خاصة في كل اللغات، جاء اقتراح التعامل معها بصورة استثنائية.

٧- لعلاج المشكلة الخاصة بالأسماء العربية والإسلامية، أنشأت الشركة المنفذة لمشروع التعريب قاعدة معلومات خاصة بها، بها حوالي أربعين ألف اسم مع مقابلاتها المعربية، على أن يبدأ نظام التعريب من أية لغة بها قبل تطبيق القواعد الخاصة بكل لغة.

٨- جاء التوجه إلى اعتبار النظام المبني على الأسلوب الإحصائي النظام العام default في غياب أسلوب محدد للتعريب، مثلا عند عدم معرفة اللغة التي ينتمي اسم ما إليها.

٦- أساليب المراجعة:

١- كان هناك تنسيق تام بين مركز المعلومات ومستشاره اللغوي وبين الشركة المنفذة للمشروع، من حيث الاتفاق على القضايا والأسس العامة، إضافة إلى المناقشات الدورية بين الأطراف المعنية، مما يسر سير العمل في المشروع.

٢- قام المستشار اللغوي للمشروع باختبار كل نظام يتم تطويره لبلد أو لغة ما بطريقة منتظمة وشاملة، وذلك من خلال الأساليب التالية:

أ- فحص الأسماء التي يشك وجود مشكلات بها للتثبت من معالجتها بصورة صحيحة.

ب- مراجعة الأسماء التي تم تعريبها ومقابلاتها، اسما اسما للتأكد من صحة التعريب ولمراجعة الأخطاء الواردة فيها. وتجدر الإشارة هنا إلى أنه نظرا لعدم توفير الشركة المنفذة قوائم المقابلات العربية (من حيث الحروف) كان على مستشار المشروع استخلاص ذلك من فحص الأسماء المعربة.

ت- تقديم الملاحظات على ما تم فحصه من القوائم لمنفذ المشروع، وذلك لإجراء التعديلات والتصويبات اللازمة والتي يكون بعض منها خطأ من المترجمين وبعض أخرنتيجة لخطأ في تخزين الأسماء، مع مناقشة بعض القضايا الخلافية القليلة للتفاهم حيالها.

ث- إجراء فحص للحالات التي تم إجراء تعديل أو تصويب فيها، للتثبت من قيام الشركة بها.
ج- إجراء فحص عشوائي منتظم للأسماء المعربة (مثلا: اسما من كل مائة أو ألف اسم من القائمة).
ح- إجراء اختبار عشوائي للنظام المطور للتعرف على مدى دقته في التعريب، وتحديد نسبة الخطأ فيه.

٦-١- أهم أسس التقويم والمراجعة:

١- كانت القاعدة الأولى مدى سلامة التعريب، إو إيجاد المقابلات العربية للأسماء. وقد تم ذلك في ضوء (أ) المعلومات اللغوية العلمية لنظامي الكتابة والنطق في اللغة المعنية، ويشمل ذلك نظام الرومنة للغات التي لاتستخدم الحرف اللاتيني، كالروسية واليابانية والصينية واليونانية. (ب) التحليل اللغوي للقائمة المعربة. (جـ) الرجوع إلى بعض أهل اللغة واستنتاج المعلومات منهم أواستشارتهم إن كانوا أهلا لذلك، وقد شمل ذلك الرجوع أحيانا إلى بعض سفارات الدول الأجنبية في الرياض. (د) التجربة والمعلومات الشخصية للفاحص.

٢- أما القاعدة الثانية فكانت الاطراد consistence أو غياب التناقض والاختلاف في التعامل مع الأسماء المتشابهة في أسماء البلد الواحد مثلا.

الملحق (١)

أساليب استخدام الحرف اللاتيني

١- الأساليب:

هناك عدة أساليب تستخدم بها الحروف اللاتينية، منها: صوامت (حروف ساكنة) مفردة ومركبة من صوامت مع صوامت ومركبة من صوامت مع صوائت (حركات). ومثال الأسلوب الأول ما نجده في اسم مثل Mark، حيث نجد أن كل صامت ينطق بطريقة واحدة (م ر ك). أما الأسلوب الثاني فنجده في اسم مثل Philips، حيث نجد أن الحرفين الأول والثاني يمثلان صوتا واحدا هو الفاء. أما المثال على الأسلوب الثالث فنجده في أسماء مثل Davenci، حيث نجد أن حرف c ينطق (تش) عندما يليه صائت (حركة) أمامية مرتفعة أو متوسطة (i,e). هذا عن الصوامت. أما الصوائت فهي أيضا تستخدم منفردة، كما في Anthony (أنتوني)، حيث تنطق كل حركة (صائت) بصورة مستقلة (ا، و، ي). كما أنها تستخدم مصحوبة بصوائت أخرى، كما نجد في استخدام ee, ie, ei للتمثيل للكسرة الطويلة، كما في اسم Keith و Green و au, eu التي تنطق ضمة مائلة في الفرنسية مثلا.

وبجدر بالذكرأن الحروف اللاتينية قد خضع بعضها لإضافات عليها للتمثيل لأصوات معينة، كما فعلت الفرنسية بإضافة s صغيرة تحت حرف c (وهو مايسمى سيديا ç) لتدل على أن الحرف المذكور ينطق سينا. وفي بعض اللغات الرومانية الأصل نجد إضافة علامة على الحرف n لينطق (ني ň). وينطبق الشيء نفسه على الحركات، حيث نجد الأمثلة التالية: ä، ō، إضافة إلى علامات النبر المختلفة التي تضاف في بعض اللغات، مثل الفرنسية.

تيسيرا على القارئ سنحاول التركيز على الحروف اللاتينية الأصلية، وذكر مقابلاتها بالحرف العربي. ثم نناقش بعض المشكلات التي تعترض كتابة الأسماء المكتوبة بالحرف اللاتيني بالحرف العربي. وسنستخدم الترتيب المتبع في الإنجليزية لمعرفة أكثر الناس بها.

الجدول (١) الصوامت (الحروف الساكنة) الأصلية:

الحرف اللاتيني الصوت	بالحرف العربي المقابل نظريا	المقابل عمليا

B	ب	ب	ب
C	ك< س/تش>	ك	ك /حسب النطق
D	د	د	د
F	ف	ف	ف
G	گ <ج/خ>	گ	ق/غ/ج/ك
H	هـ	هـ	هـ
J	ج/ي/خ/هـ	ج	ج /حسب النطق
K	ك	ك	ك
L	ل	ل	ل
M	م	م	م
N	ن	ن	ن
P	پ	ب	ب/پ (المثلثة)
Q(U)	ك	ك	ك
R	ر	ر	ر
S	س /ز	س/ ز	س /ز
T	ت	ت	ت
V	ف/ ڤ (مثلثة)	ف	ف/ ڤ (مثلثة)
W	و/ ڤ (مثلثة)	و	و
X	ك+س / ز	ك+س / ز	ك+س / ز
Y	ي	ي	ي
Z	ز/تز	ز	ز

نلاحظ في الجدول أعلاه أننا أوردنا المقابلات المشهورة للحرف اللاتيني الساكن. ومع ذلك نجد أن بعض هذه الحروف تنطق بأشكال مختلفة تبعا للغة الأوربية التي تستعمل فيها أو للسياق (موقعها في الكلمة أو الصائت الذي يرد بعدها مثلا)، كما نوضح أدناه.

أ- تأثير اللغة:

لعل من أبرز أمثلة تأثير اللغة، أو حتى ا للهجة من لغة ما، الحروف التالية:

J: فهذه تنطق بأربعة طرق مختلفة هي: (١)/ج/ في الإنجليزية و/چ/ معطشة في الفرنسية، (٢) /ي/ في الألمانية واللغات الاسكندنافية (بل وفي اللاتينية الأصلية)، (٣) /خ/ في الأسبانية الأوربية، (٤) /هـ/ في أسبانية أمريكا الجنوبية والمكسيك.

S:/ س/ في معظم اللغات (إذا استبعدنا تأثير السياق)،/ ز/ في الألمانية (كما في Siemens (زيمِنز).

V: (١)/ ف/ ڤ مثلثة (بثلاث نقاط فوقها) في معظم اللغات الأوربية (٢) و/ ف/ الألمانية (كما في فولكسVolks).

Z: (١)/ ز/ في معظم اللغات الأوربية (٢)/ تز/ في الألمانية والإيطالية.

ب- أمثلة على تأثير السياق الحرفي:

لعل من أبرز أمثلة هذا النوع من السياق ما يلي:

C : (١) /ك/ في اللاتينة الأصلية وحينما لاتليها حركة (أو صائت) مرتفع، مثل الكسرة أو الكسرة المائلة وما في حكمها (y, e, i). (٢) /س/ حينما تليها كسرة أو ما في حكمها (y, e,i). (٣) /ت+ش/ في اللغة الإيطالية حينما تليه حركة مرتفعة (i,e)، كما في العالم والفنان Davenci (دافنتشي).

G: (١) /گ/ (وهو ما يشبه الجيم القاهرية والقاف السعودية العامية) في اللاتينة الأصلية وحينما لاتليها حركة (أو صائت) مرتفع، مثل الكسرة أو الكسرة المائلة وما في حكمها (y, e,i) في معظم لغات أوربا الغربية (٢) /ج/ حينما تليها كسرة أو ما في حكمها (y, e,i)، مع ملاحظة أن الأسبانية في لهجتيها الرئيستين ليس بها صوت /ج/. فهي تستخدم /خ/ أو /هـ/ بدلا من /ج)، كما هو الحال مع الحرف J.

X: (١) في معظم الأحيان تنطق/ ك+س/ أو/ك+ز/، كما في Alex (أَلِكس) أو Alexander (أَلِكزاندَر). (٢) أما إذا وردت في أول الكلمة (غالبا في كلمات مقترضة من اليونانية) فتنطق /ز/، كما في (زينون Xenon). ومن الأمثلة التي يرد فيها النطقان معا اسم Xerox (زيروكس). هذا وقد وضعنا المتغيرات الصوتية تبعا للسياق بين قوسين مثلثين < > في العمود الثاني أعلاه.

ج- تأثير السياق الصوتي:

هناك أيضا تأثير السياق الصوتي للحرف، فحرف s مثلا، على الرغم من مقابلته السين العربية في معظم اللغات، إلا أنه ينطق زايا في بعض السياقات الصوتية كأن يسبقها صائت ويتبعها صوت مجهور، كما في اللاحقة ism– التي تدل على المصدرية في بعض اللغات الغربية.

د- أثر القواعد الإملائية/ الصوتية:

يجدر بنا أن نشير إلى أثر مهم في نطق بعض الحروف أوحتى عدم نطقها تبعا لقواعد الإملاء أو الأصوات في لغة ما. ففي الإنجليزية مثلا لاتنطق k إذا تبعها حرف ساكن في أول الكلمة. كذلك نجد أن g لاتنطق إذا جاءت في آخر الكلمة متبوعة ب n، كما في sign (ساين) (يلاحظ هنا أن الصائت i ينطق بطريقة مختلفة عنها في الكلمة التي لاتكون فيها gn في آخر الكلمة، كما في signature (سِگنِتشَر)). وفي الفرنسية نجد أن الحرف الساكن لاينطق غالبا إذا ماورد في آخر الكلمة، كما في الاسم Robert (روبير)، حيث لاتنطق التاء الأخيرة. وأخيرا نجد أن كثيرا من لغات أوربا الغربية – إن لم تكن جميعها – مما تستخدم حرف q تأتي دائما متبوعة بالحرف u، مع اختلاف في النطق: فالإنجليزية والأسبانية والإيطالية مثلا تنطق الحرفين كصوتين /ك+و/، كما في quick (كوِك)، بينما تتجاهل الفرنسية الصائت u تماما، كما في Jaque (جاك).

الجدول (٢) أمثلة للأصوات التي تمثلها مجموعات مركبة من حرفين ساكنين أو أكثر:

المجموعة	الصوت	النقل العربي لها
Ch	تش/ش/خ/ك	حسب نطقها/ ج
Ph	ف	ف
Sh	ش	ش
Sch	ش/سك	حسب نطقها
Th	ث/ذ/ت	حسب نطقها

يلاحظ هنا أن عامل اللغة التي ترد فيها المجموعة لها أثرها. فالمجموعة ch كما نرى أعلاه تنطق عادة (١) /تش/ في الإنجليزية و(٢) /ش/ في الفرنسية و(٣) /ك/ في لغات لاتينية الأصل أخرى، مثل الإيطالية. (٤) أما في الألمانية فإن النطق يعتمد على الصائت (الحركة) الذي يأتي قبل ch، فإذا كانت ضمة أو فتحة مثلا فإن المجموعة تنطق /خ/، كما في Bach (الموسيقار باخ). وتنطق بين الشين والخاء إذا وردت بعد الصائت المرتفع i كما في Ich (إيش/خ).

أما المجموعة gn فإنها تنطق ن+ي في اللغات ذات الأصل اللاتيني عادة، كما في signor (سنيور). أما في الإنجليزية فالحرفان إما أن ينطقان صوتين مختلفين أو لاينطق أحدهما، كما أشرنا أعلاه.

بالنسبة للمجموعة sh فهي شائعة في الإنجليزية لتمثل الصوت /ش/. وتقابلها في الفرنسية ch. أما في الألمانية فتكتب sch. ويجدر بنا أن نشير إلى أن sch تنطق في الإنجليزية /سك/، كما في school (سكول) و في النطق الأمريكي لكلمة schedule (سكِدِيُل)، أو /ش/، كما في النطق البريطاني لكلمة schedule (شِدِيُل).

أما المجموعة th فإنها تخضع في نطقها لكل من عاملي اللغة والسياق اللفظي. فهي مجرد /ت/ في الفرنسية مثلا، كما في theorie. أما في الإنجليزية فتنطق في ثلاث صور مختلفة: (١) /ت/ في بعض أسماء الأعلام مثل Thomas (توماس) وAnthony (أنتوني – عربت إلى أنطونيو، بناء على الاسم الروماني). (٢) /ث/ كما في Thebe (ثيبي). (٣) /ذ/ كما في this (ذِس).

الجدول (٣) أمثلة لثنائيات من الصوامت تتغير صفة مكوناتها أو بعضا منها:

الثنائية	نطقها	أمثلة
Gn	ني	signor (سِنيور)
St	شت	Stein (شتاين)

الجدول (٤) أمثلة للأصوات التي تمثلها مجموعات مركبة من حرف ساكن وصائت:

المجموعة	الصوت	النقل العربي لها
Ci	ش/تش	حسب اللغة والنطق
Si	چ/زي	ج/زي
Su	چ/ش	ج/ش
Ti	ش /سي	ش/ سي
Tu	تش	تش
Ill	إي	إيل

نجد هنا أن كلا من عاملي اللغة والسياق اللفظي له دوره في التلفظ بهذه المجموعات من الحروف. ففي حالة ci مثلا نجدأنها تنطق /ش/ في الإنجليزية فقط، وفي بعض السياقات، كما في كلمات مثل special وassociate (أسوشييت) في النطق الأمريكي، علما بأنها تنطق /ش+ي/ في هذه الأخيرة وليست مجرد /ش/. وفي الإيطالية ينطق هذا الثنائي ومثله ce/تش/، كما في ciao (تشاو). أما في أكثر اللغات الأوربية الأخرى فلا تعامل ci كوحدة واحدة – بل هي مجرد /س+ي/، كما في كلمة sociale الفرنسية (سوسيال). وينطبق ذلك حتى على نطق بعض الكلمات في الإنجليزية البريطانية، كما في كلمة associate المذكورة أعلاه والتي تنطق (أسوسييت).

والشيء نفسه ينطبق على si، فهي تنطق جيما معطشة في الإنجليزية، كما في vision (فِچَن) (ولكن ssio تنطق /ش/، كما في mission (مِشَن)). أما في الفرنسية فهي /ز+ي/، كما في نطق الكلمة نفسها (فِزيون).

ونجد الحالة نفسها مع su في سياقات معينة في الإنجليزية، كما في measure (مِچَر)، وهي غالبا ما تكون /ز+و/ في الفرنسية. غيرأن هناك نطقا آخر لهذه المجموعة من الحروف في الإنجليزية، وهي /ش/ كما في sure (شور).

والإنجليزية ريما تنفرد كذلك في النطق الخاص ب ti, tu بوصفهما /ش/ أو /تش/ في بعض السياقات، كما في nation (نيشَن) و nature (نيتشَر)، حيت نجد ti في الكلمة المذكورة سابقا تنطق /س+ي/ في الفرنسية،فتنطق الكلمة نفسها (ناسيون).

أما tu فنجدها في الفرنسية /ت+و/ (ناتور)، مع ملاحظة النطق الخاص بالحرف u في الفرنسية (كسرة مضمومة).

وهناك تفسير علمي صوتي لمثل هذه الاختلافات في نطق هذه المجموعات، ألا وهو مايسمى بالمماثلة الصوتية، حيث يؤثر صوت ما في نطق صوت مجاور آخر، كما نجد في حالات الإدغام والإخفاء المختلفة في اللغة العربية. ونجد أن المماثلة قد تحدث في لغة أو سياق معين ولا تحدث في لغة أو سياق آخر.

المثال الأخير أعلاه يعتبر سائدا في اللغات ذات الأصل اللاتيني، كالفرنسية والأسبانية حيث نجد أن ill قد تنطق /إي/ (أي يتحول اللام إلى ياء مشددة) في بعض السياقات الصوتية أو الإملائية، كما في الكلمة الفرنسيةfille (فِيّ). ومن أمثلة هذا التحول اسم المدينة الفرنسية Marseille مارسيليا، حيث أبقى النطق العربي على اللام الموجودة في رسم الكلمة، بينما تنطق مارسِي في الفرنسية.

الصوائت (الحركات):

الجدول (٥) الصوائت الأساسية:

الحرف اللاتيني	الصوت
A	فتحة طويلة/قصيرة / كسرة مائلة طويلة
E	كسرة مائلة طويلة/قصيرة/ضمة
I	كسرة قصيرة/ طويلة/ آي
O	ضمة مائلة قصيرة/ طويلة/ ا مفخمة
U	ضمة قصيرة/ طويلة
	/كسرة مضمومة/ يو/ فتحة قصيرة
(Y)	كسرة قصيرة/ آي

نلاحظ هنا مايلي:

١- يلاحظ أن الخيار الأول هو الأشيع، وتليه الخيارات الأخرى.

٢- استعمال الa للدلالة على الكسرة الطويلة المائلة في بعض الألفاظ الإنجليزية، مثل Jane (جين) تنفرد بها اللغة الإنجليزية، وربما يرجع ذلك إلى أن اسم الحرف جاء بهذه الصورة، على عكس اسمه في اللغات الأخرى التي تنطقها /آ/ أي فتحة طويلة. والكلام نفسه ينطبق على نطق الصائت i، حيث إنه يتأثر باسم الحرف كذلك، كما في آيس .ice

٣- تنطبق ظاهرة تأثيراسم الحرف على نطقه في حالة u في الإنجليزية كذلك، حيث نجده ينطق /يو/ في كلمات مثل Duke. فالحرف يسمى حرف اليو، بخلاف مانجده في اللغات الأوربية الأخرى.

٤- تنفرد الإنجليزية بنطق ال u فتحة قصيرة، كما في اسم الشاعر الأمريكي Cummings (كَمِنقز).

٥- النطق بحرف e على شكل ضمة مائلة قصيرة لا نجده سوى في الفرنسية، حسب علمنا. فمثلا تسمى كندا Le Canada (لو كَنَدا). لذلك كثيرا مانجدها تعرب إلى كسرة (مائلة -لي) بحكم نطقها السائد في اللغات الأخرى، والشيء نفسه نجده في اسم الرئيس الفرنسي الراحل De Gaule (ديقول).

٦- تنفرد الفرنسية أيضا بنطق ال u على شكل كسرة مضمومة، كما في tune (تِيُن)، ولكن الغالب في نقحرتها هوالواو بسبب شكل الحرف غالبا. أما ما يقابل الضمة العربية القصيرة والطويلة فإنها تكتب ou، كما في الكلمة الفرنسية mouchoire (مُشوار).

٧- النطق بالحرف بوصفه حركة طويلة أو قصيرة يخضع غالبا لقواعد صوتية، مثل وجود النبر القوي من عدمه على سبيل المثال.

٨- هناك قواعد لغوية تغير من نطق الحركة تماما، كما نجد في تحويل a من فتحة إلى ضمة طويلة في الكلمات الإنجليزية التي تنتهي بحرفي ll (اللام) مسبوقة بهذه الحركة، كما في mall, call وغيرها.

٩- على الرغم من أن الحرف y حرف ساكن، إلا أننا نجده يمثل حركة الكسرة أحيانا أو ينطق /آي/ ، كما في synonym (سِنونِم) و Kyle (كايل).

١٠- تتميز الإنجليزية بتأثير الراء r على بعض الصوائت، خاصة i, u و تغير طبيعتها، فتجعلها كسرة طويلة مائلة، كما في curtain (كيرتِن) وsir (سير).

١٠- يلاحظ أن الإنجليزية الأمريكية تنفرد بنطق ال o فتحة طويلة ألف مد مفخمة قليلا، كما في shock, hot (شاك، هات) مما يجعل الكلمتين تختلطان بالكلمتين shark, Hart في الإنجليزية البريطانية الرسمية والتي تحذف الراء في الحالتين المذكورتين وتفخم الألف.

١١- تتميز الفرنسية بخاصية غنة الحركات قبل حرف n، مما يجعلها تخضع للتغيير في صفتها، كما في fin التي تنطق مثل (فَن – بنون خفيفة)، وليس (فين)، كما هو المتوقع.

١٢- في الفرنسية نلاحظ أيضا أن الحرف e (الذي ينطق ضمة مائلة قصيرة عادة) يصبح كسرة مائلة طويلة إذا ماتلاه r، كما في Robert (روبير) أو t كما في Roujet (روجيه).

١٣- الصائت e لاينطق إذا ماورد في آخر الكلمة الإنجليزية، مالم تكن الكلمة ذات أصل أجنبي، كما في Monte Carlo، حيث تنطق الكلمة الأولى (مونتي).

الجدول (٦) أمثلة لمجموعات الصوائت التي تنطق حركات واحدة:

الحرف اللاتيني	الصوت	المثال
AY	كسرة طويلة مائلة	say
AI	كسرة طويلة مائلة	bait
AU	ضمة طويلة مائلة	sauvage, Paul
AW	ضمة طويلة مائلة	saw
EA	كسرة طويلة (ي)	seat
EAU	ضمة مائلة طويلة	beau
EE	كسرة طويلة (ي)	feet
EI	كسرة طويلة (ي)	receive
EU	ضمة مائلة طويلة	jeune
EY	كسرة طويلة (ي)	key
IE	كسرة طويلة (ي)	sieve
OA	ضمة مائلة طويلة	boat
OO	ضمة طويلة / قصيرة	book, moon
OU	ضمة طويلة/ فتحة قصيرة	jour, colour
OW	ضمة طويلة مائلة	know

ملاحظات:

١- يلاحظ أننا أوردنا أهم المجموعات التي تعامل كل مجموعة منها حركة واحدة ، علما بأن بعض المجموعات قد تنطق حركة واحدة أحيانا وحركتان أحيانا أخرى في نفس اللغة، كما في حالة ow التي تنطق في الإنجليزية ضمة طويلة،كما في sow (سو) أو حركتين، كما في now (ناو). كذلك يلاحظ أننا أوردنا شبها الصائتين w, y حيث إنهما تعاملان كصائتين أحيانا إذا جاءتا بعد بعض الصوائت (الحركات).

٢- يأتي تكرار الحرف للدلالة على الطول في بعض اللغات، كما نجد في الاسم الإسكندنافي Aartz.أما في الإنجليزية فإن التكرار غالبا مايغير من صفة الصائت، وهو مانجده في المثالين ee, oo، حيث إن ee ليست كسرة مائلة طويلة كما نتوقع، بل هي كسرة طويلة غير مائلة، كما في seek (سيك). وال oo ليست ضمة مائلة طويلة، بل تنطق إما ضمة قصيرة، كما في book (بُك) أو ضمة طويلة غير مائلة، كما في Booth (بوث).

٣- يمكننا أن نجد أمثلة على جميع المجموعات المذكورة بنطقها المذكور أمام كل منها في اللغة الإنجليزية، باستثناءين، هما ou التي ترد صوتا واحدا، هو الضمة الطويلة غالبا والقصيرة أحيانا، في الفرنسية، كما في journal (جورنال) و mouchoire (مُشوار). أما في الإنجليزية فتنطق غالبا حركتان، كما في out (آوت)، ولو أننا نجدها تنطق ضمة كما في بعض الكلمات ذات الأصل الفرنسي، مثل moustache. والاستثناء الثاني هو المجموعةالمكونة من ثلاثة حركات eau التي ترد في الفرنسية وتنطق ضمة مائلة طويلة، كما في beau (بو).

٤- في الإنجليزية نجد حالات استثنائية ترد فيها ae وتنطق كسرة طويلة، كما في اسم الامبراطور الروماني Caesar (تنطق سيزَر) وجموع بعض الكلمات المقترضة من اللاتينية، مثل formulae التي تنطق (فورميولاي أو فورميولي).

الجدول (٧) مجموعات الصوائت التي تنطق مختلفة:

هناك عدد من الصوائت (وأشباه الصوائت) التي ترد متجاورة ولكنها تنطق مستقلة عن أحدها الآخر، وسنورد أدناه بعضا من أكثرها شيوعا:

الحرف اللاتيني	الصوت	المثال
AE	آءِ	rosae
AO	آو	ciao
AU	آو	Raul
AYE(R)	كسرة مائلة طويلة وياء وفتحة	player
EI	آي	Heiden
EO	إيو	Leo
IO	آيو	Lions
OU	آو	out
OW	آو	now
OI	ضمة مائلة طويلة وياء/ وا	boil, soir
OY	ضمة مائلة طويلة وياء/ وا	Joyce

ملاحظات:

١- نجد المثال الأول بكثرة في اللاتينية، حيث تأتي صيغة الجمع للكلمات المنتهية ب a بزيادة e بعدها. وتحتفظ بعض اللغات الأوربية بهذا النطق عند استعارة ألفاظ لاتينية، مثل formula وجمعها formulae.

٢- المثالان الثاني والثالث يردان بكثرة في بعض اللغات ذات الأصل اللاتيني، كالأسبانية والإيطالية، كما في المثالين المذكورين أعلاه: ciao (تشاو) Raul (راؤل).

٣- المثال على المجموعة الثالثة نجده في الإنجليزية خاصة عند إضافة اللاحقة er– على الكلمات لتجعلها أسماء، مثل player، أو على الصفات للاتيان بصيغة المقارنة، كما في gayer.

٤- نجد النطق /آي/ لهذا الثنائي من الصوائت في الألمانية، كما في اسم المدينة Mein (ماين) والموسيقار Heiden (هايدِن). وننبه إلى أن هذين الصائتين ينطقان كسرة طويلة في الإنجليزية، كما ذكرنا سابقا.

٥- الثنائي eo يرد في كثير من اللغات، كما في اسم برج الأسد Leo.

٦- لانجد هذا النطق ل io حسب علمنا إلا في الإنجليزية، كما في Lions (لايونز).

٧- الثنائي ow نجده في الإنجليزية، كما في how (هاو). وقد أشرنا سابقا إلى أن المجموعة نفسها قد تنطق ضمة مائلة طويلة، كما في low.

٦- الثنائي io ينطق بطريقتين مختلفتين، حسب اللغة. فهو في الإنجليزية /أُوي/، كما في coil (كُويل) بينما ينطق /وا/ في الفرنسية، كما في soir (سوار).

٧- الحرفان oy ينطقان في الإنجليزية مثل oi، كما في اسم الروائيJoyce (جُويس). كذلك الحال في الفرنسية، حيث تعامل oy مثل io، كما في foyer (فوابيه).

٢- أهم مشكلات كتابة الأسماء الغربية بالحرف العربي:

هناك عدد من المشكلات التي تواجهنا عند نقل الأسماء المكتوبة بالحرف اللاتيني إلى الحرف العربي، تشمل مايلي:

١- اختلاف نطق بعض الحروف الساكنة والحركات (الصوامت والصوائت) تبعا للغات والسياقات والقواعد اللغوية (الإملائية والصوتية)، حتى في لغات أوربا الغربية ذات الأصول الجرمانية أواللاتينية.

٢- الاختلاف البين في سلوك لغات بعض أوربا الشرقية (مثل التشيكية والصربية والبولندية)عند استخدام بعض الحروف اللاتينية وفي نطقها. وهو مالم نتعرض له في بحثنا هذا.

٣- مشكلة الأسماء ذات الأصول العربية، والتي تتطلب معرفة بأصولها عند كتابتها بالحرف العربي. ولعلنا لاحظنا اختلاف وسائل الإعلام في كتابة اسم الرئيس الأفغاني (هل هو "كرزاي" أم "قرضاي"؟ وفي كتابة اسم رئيس وزراء ماليزيا السابق Mahazir (هل هو "محاذير/ محازير" أم "محاضير" أم "محاضِر"؟ واسم وزير خارجية إيران (التي ينقلها بعضهم "متكي" والآخرون "متقي") وهكذا دواليك. إضافة إلى أن بعض اللغات الإسلامية التي تمت رومنة كتابتها تستخدم بعض الحروف اللاتينية بصورة ونطق مختلفين عما تستخدم فيها في اللغات الأوربيةالغربية.

تجنبًا للإطالة فإننا سنركز حديثنا على مشكلات كتابة الأسماء في اللغات الأوربية الغربية، أي المشكلات المشار إليها في المجموعة الأولى من المشكلات، مؤجلين الحديث عن المشكلات في الفئتين الثانية والثالثة إلى مناسبة أخرى إن شاء الله

٤- أهم مشكلات كتابة الأسماء الواردة في لغات أوربا الغربية:

من قراءة متأنية لجداول الحروف الساكنة والمتحركة أعلاه، يمكننا ملاحظة مايلي:

أ- على الرغم من أن جميع هذه اللغات تستعمل الحرف اللاتيني إلا أن هناك اختلافات في نطق بعض الحروف سواء أوردت منفردة (مثل j) أم مجتمعة مع حروف أخرى (مثل: th, st, gn, ci, ti) وغيرها مما أوردنا أعلاه. والمشكلة أكبر مع الصوائت (الحركات). فكما نلاحظ من جدول الصوائت الأساسية (الجدول ٥)، أنه حتى هذه لم تسلم من الاختلافات في نطقها، وليس من صائت واحد له نطق واحد في جميع لغات أوربا الغربية، بل وفي اللغة الواحدة أحيانا كما في الإنجليزية.

ب- كما ذكرنا أعلاه، يختلف نطق هذه الحروف ليس فقط تبعا للغة التي تستخدم فيها، بل وتتأثر كذلك بعوامل أخرى مثل الحروف التي قبلها والتي بعدها، كما نرى من جداول ثنائيات أو مجموعات الحروف الساكنة والمتحركة أعلاه (الجداول ٢و٣ و٤ و٦ و٧). كما إنها تخضع لقواعد إملائية وصوتية كذلك، مثل منع الإنجليزية ورود صوت انفجاري (مثل k, p) قبل ساكن آخر في بداية الكلمات، من ثم يصبح هذا الصوت "ساكتا" – لاينطق به، كما في know (نُو)، بينما تجيز لغات أخرى كالألمانية مثل هذه الظاهرة، ومثل عدم نطق الإنجليزية لحرف e عندما ترد بعد حرف ساكن في آخر الكلمة وعدم نطق الفرنسية للحرف الساكن عندما يرد في آخر الكلمة أحيانا إلى غير ذلك من الأمثلة التي لاحصر لها.

فلا غرو أنه حتى أهل اللغة يعجزون في كثير من الأحيان عن كتابة الأسماء غير المألوفة بصورة صحيحة. من ثم كثيرا مانسمع في البلدان الناطقة بالإنجليزية السؤال المعروف: "How do you spell it? How do you spell your name?" أي "كيف تتهجاه؟ أو كيف تتهجى اسمك؟"

جـ- هناك مشكلات لغوية تتعلق بنظام الكتابة العربية، من أهمها (١) عدم وجود مقابلات عربية لبعض الحروف اللاتينية، مثل p, v, g على الرغم من وجود مقابلات لها، مثل الباء والفاء المثلثتين (بثلاث نقط) والكاف الفارسية، مما تستعمله بعض اللغات الإسلامية، ومثل الحركات المائلة (الكسرة والضمة المائلتين، كما في Jane, John مثلا) والتي لارموز لها في العربية أو اللغات التي تكتب بالحرف العربي (حسب علمنا)، مما يؤدي إلى اختلافات أحيانا في نقل بعضها إلى العربية، كما في كتابة الاسم Jane (هل يكتب جاين أو جين؟).

(٢) عدم استعمال الحركات القصيرة (الفتحة والضمة والكسرة) في الكتابة العربية المعتادة، على الرغم من وجودها في اللغة العربية، مما يلجيء كثيرًا من معربي الأسماء الأجنبية إلى إضافة ألف أو واو أو ياء للتعويض عنها عند كتاب الأسماء الأجنبية أحيانا، مثل كتابة الاسم Bush في صورة (بوش، وليس بُش كما تُنطق) و Munich على شكل (ميونيخ، بدلا من ميونِخ) وهكذا.

(٣) اختلاف العرب في كتابة بعض الحروف اللاتينية التي لها مقابل عربي، مثل حرف t التي تكتب تاء وطاء أحيانا (طوني وطاكسي بدلا من توني وتاكسي)، علما بأن بعض العرب الأوائل فعلوا ذلك، كما في أرسطوطاليس وسقراط.

كذلك نجد أن صوت الكاف (التي تكتب أحيانا c بالحرف اللاتيني) تم تعريبها إلى ق، كما في موسيقى و سقراط (مقابلاتهما في الإنجليزية music, Socrates).

والأمر أسوء في حالة الحروف التي لايوجد لها مقابل في العربية، مثل g، كما في Gulliver فالشاميون يكتبونها /غ/ والمصريون يكتبونها /ج/ والعراقيون يكتبونها /ك/ وبعضهم في دول الجزيرة العربية يكتبها /ق/، كما في: اللغة الإنغليزية، الإنجليزية، الإنكليزية، الإنقليزية.

جدير بالذكر أن العرب الأوائل كانوا يعربون الحرف إلى غ، كما في فيثاغورس.

د- إضافة إلى ماورد أعلاه، هناك مشكلات أخرى تكمن في ضرورة لجوء المترجم العربي إلى معرفة متى يعرب الكلمة (يكتبها بالحرف العربي) ومتى يورد المقابل العربي لاسم العلم الذي يتعامل معه. فاسم العلم Toledo على سبيل المثال ربما يجب أن يكتب (توليدو) إذا كان يشير إلى المدينة الأمريكية في ولاية أوهايو. ولكن لا بد لنا أن نورد مقابلها العربي إذا كنا نتحدث عن المدينة الأسبانية المعروفة بطليطلة. وكذا الحال مع اسم Alexandria إذا وردت في سياق مدينة أو قرية في دولة غربية فإننا نبقي على النطق (ألكزاندريا). أما إذا كانت الإشارة إلى اسم المدينة المصرية المعروفة فلا بد أن نعطي المقابل (الإسكندرية). بل وينطبق هذا على أسماء الدول التي لها مقابلات عربية. فكلمة Belgium مثلا عندما ترد في نص إنجليزي لاتعرب ولكن يعطى المقابل العربي المعروف (بلجيكا) والأمر كذلك مع كثير من أسماء الأعلام. فالاسم Aristotle في الإنجليزية يعطى مقابلة العربي (أرسطو) واسم الفيلسوف العربي الذي يكتب Averroes لايكتب (أفيروس)، بل ابن رشد.

الملحق (٢)

مقترحات بخصوص البلدان التي وردت أسماؤها في القائمة الأخيرة لدى الجهة صاحبة المشروع

البلدان مرتبة ألفبائيا.

البلد	اللغة المقترحة لها أوالبلد المقترح دمجها معها
إثيوبيا	الأمهرية
ألبانيا	الألبانية
أوزبكستان	مع تركيا
أيسلندا	Generic
بخارستان	مع تركيا
بلغاريا	البلغارية
بهوتان	Generic
البوسنة والهرسك	مع يوغسلافيا
تايلاند	التايلاندية
تركستان	مع تركيا
تركمانستان	مع تركيا
تشيك	Generic
جرينلاند	مع الدانمارك
جزر المالديف	Generic
جزر فيرو	مع الدانمارك

جوانا (ربما Guyana)	Generic
الدانمارك	الدانماركية
سلوفاكيا	مع يوغسلافيا
سلوفينيا	مع يوغسلافيا
السويد	السويدية
صربيا والجبل الأسود	مع يوغسلافيا
طاجكستان	مع روسيا أو إيران
فنلندا	الفنلندية
قبرص	اليونانية
كرواتيا	مع يوغسلافيا
كوريا الجنوبية	الكورية
كوريا الشمالية	الكورية
كوسوفوا	مع يوغسلافيا
لاوس	مع فيتنام
مقدونيا	مع يوغسلافيا
منغوليا	الروسية
مهرة	Generic
ميانمار	Generic
النرويج	النرويجية
نيبال	Generic
نيو	Generic

هنغاريا	المجرية
اليابان	اليابانية
يوغسلافيا	الصربية
اليونان	اليونانية

ملحوظة هامة:

تم اقتراح القائمة بناء على دراسة نظرية وعملية (وفق المتوافر من قوائم الأسماء). ولكن يراعى عند تطبيق القواعد العمومية Generic بأن حرف e في أواخر الأسماء ليست صامتة silent كما في الإنجليزية، بل إنها منطوقة، كما أنها في جميع الأسماء الإسلامية تقريبا تمثل تاء التأنيث المربوطة.

صناعة المعاجم العامة والمتخصصة (المصطلحات)

الحاسوب في العمل المعجمي

مقدمة:

من المعروف أن المعاجم هي خزانة اللغة، فهي تحفظ ألفاظها، ومعها الكثير من تراثها الفكري.وقد كان العرب روادًا على المستوى العالمي في صناعة المعاجم ووضعها، بدءًا بكتب الغريب والموضوعات المتخصصة (كالخيل)، ثم المعاجم اللغوية العامة التي كان للخليل بن أحمد الفراهيدي قصب السبق في إعدادها بمعجمه المعروف ب (كتاب العين).

وقد لجأ الخليل إلى عدة أساليب في عمله: (١) الاستقصاء المنطقي للتعرف على الألفاظ الممكنة عن طريفق التلقليبات ثم تحديد المستعمل منها والمهمل و (٢) الترتيب المنطقي المنتظم لمداخل المعجم، مستعينًا بنطق جذورالألفاظ ومخارج أولها. فرتب المعجم معتمدًا تلك المخارج في الترتيب، حيث بدأ بالكلمات (الجذور) التي تكون مخارجها في أقصى الحلق (حرف العين) ثم متدرجًا إلى أن وصل إلى الشفتين، ثم أتبعها بالحركات الطويلة، أو ماسميت بالحروف الهوائية (أ، و، ي).
(٣) اتبع الخليل منهجًا مهمًا في التعامل مع الألفاظ التي جردها من الزوائد والمشتقات ليصل إلى الجذر. ولا شك أن ترتيب المعاجم وفق جذورها كان ولا يزال الأسلوب الأشيع في ترتيب مداخل المعاجم العربية، قديمًا وحديثًا. فهو المنهج المتبع على سبيل المثال في **(لسان العرب و الصحاح وتاج العروس nortوالقاموس المحيط)** قديمًا و**(المعجم الوسيط)** حديثًا. (انظر: نصار(١٩٨٨) وشهيد الإسلام _٢٠١٠).

١- أنواع المعاجم:

يمكننا النظر في الأنواع من عدة زوايا، بدءًا
(أ) بالجانب الشكلي: ترتيب المداخل بحسب:

١-١- الجذر (مع اختلاف يسير، كما في نظام الفصل والباب في القاموس المحيط)، وهو مثل ك، ت، ب، حيث نجد تحته مشتقاته مثل: كاتب، مكتوب، كتابة...

١-٢- نظام التقفية ، أي ترتيب الكلمات والألفاظ بحسب أواخر حروفها بعد تجريد اللفظ من أحرف الزيادأي بحسب الجذر ولكن الترتيب على أساس آخر الكلمة (الباب) ثم أول الكلمة (االفصل)، مثلا كلمة "عين" نفتح على باب النون ثم فصل العين ثم الياء. وهوالمنهج الفريد الذي اتبعه الفيروزبادي في معجمه (القاموس المحيط). ومن الطريف أن نجد صدى لهذا الترتيب وفق أخر الكلمة في بعض المعاجم الإنجليزية المسماة بمعجم القافية rhyming dictionary.

١-٣- نطق الكلمة في أبسط صورها بعد تجريدها من الإضافات التصريفية: الاسم المفرد مثلا والفعل في صيغة الماضي للفاعل المذكر المفرد الغائب. وهو ما يسمى بجذع الكلمة. فنجد "كاتب" تحت الكاف والألف والتاء والباء. و"مكتوب" تحت الميم والكاف والتاء والواو والباء.

ويلاحظ أن هذا منهج كان متبعًا في معاجم المصطلحات التراثية، مثل (**التعريفات**) للجرجاني و(**كشاف اصطلاحات الفنون**) للتهانوي و(**مفاتيح العلوم**) للخوارزمي. ويشيع هذا الأسلوب في الوقت الحاضر في المعاجم الموجهة للطلاب العرب مثل (**رائد الطلاب**) لجبران مسعود و (**المعجم العربي الميسّر**) لأحمد بدوي وصِدّيقة محمود و(**المعجم العربي للطلاب**) لمجموعة من الباحثين السعوديين، أوموجهة لمتعلمي اللغة العربية من غير أهلها مثل (**المعجم السياقي للطالب**) لمحمود صيني وحيمور يوسف.

١-٤- الموضوعات: وهذه تندرج تحت ما يسمى بمعاجم المعاني. وقد كانت بدايات المعاجم العربية بموضوعات خاصة، مثل (**الخيل**) لأبي عبيد. ولكن اتضحت معالم معاجم المعاني أو الموضوعات الشاملة في أعمال رائدة أخرى، مثل (**فقه اللغة وسر العربية**) للثعالبي، و(**أدب الكاتب**) لابن قتيبة، و(الألفاظ الكتابية) للهمذاني، والعمل الجبار(**المفصّل**) لابن سيده. في العصر الحديث نجد (نجعة الوارد في المترادف **والمتوارد**) لليازجي و عددًا مما سمي بالمكانز، مثل أمثال أعمال حنا غالب وأحمد مختار عمر ومحمود الصيني وزميليه. ويدخل تحت هذه المجموعة معاجم المترادفات، مثل (قاموس المترادفات والمتجانسات) لرفائيل نخلة والأضداد كذلك. ولعل أبرز هذه المكانز (**المكنز الكبير: معجم شامل للمجالات والمترادفات والمتضادات**) لأحمد مختارعمر (مع فريق من الباحثين).

(ب) يمكن تصنيف معاجم الألفاظ إلى فئتين رئيستين: العامة والخاصة أو المتخصصة، أي المعاجم التقليدية، مثل **لسان العرب** و**المعجم الوسيط**، ومعاجم المصطلحات المتخصصة في فن أو مجموعة من التخصصات، مثل**التعريفات** للجرجاني و **مفاتيح العلوم** للخوارزمي في القديم وعشرات المعاجم المتخصصة في مجالات عديدة في العصر الحديث، مثل **المعجم الطبي الموحّد**.

(جـ) كما يمكننا أن نصنف المعاجم وفقًا للجمهور المستهدف. وهو ما نلحظه مؤخرًا في المعاجم الموجّهة للطلاب مثلًا ولمتعلمي اللغة العربية من غير أهلها، مثل **رائد الطلاب** و **المعجم العربي الميسّر** و**المعجم العربي للطلاب والمعجم السياقي للطالب** التي ذكرناها أعلاه.

- (د) وهناك تصنيف يعتمد على عدد كلمات كل مدخل في المعجم. فالمتبع هو أن يكون المدخل كلمة أو جذرًا واحدًا، مثل سائر المعاجم المذكورة أمثلتها سابقًا، ولكن هناك معاجم تتعدد فيها كلمات المدخل الواحد، كمعاجم التعابير الاصطلاحية، مثل **معجم التعابير الاصطلاحية** لوفاء فايد، و**المعجم السياقي للتعبيرات الاصطلاحية** لصيني وحسين والدوش، إضافة إلى معاجم أخرى مثل **معجم الأفعال المتعدية بحرف** لأريج المسلم، و**معجم الأفعال المتعدية بحرف لأبناء العربية الأصلاء والجدد** لعبدالقادر زينو، وكذلك معاجم المصاحبات اللفظية.

٢- الخطوات المتبعة في إعداد المعاجم:

٢-١- تحديد الهدف والجمهور المستهدف**:**

مما لاشك فيه أن تحديد الهدف يجب أن يكون أول عمل يقوم به مؤلف المعجم، له أهميته ليس في تحديد (١) نوع المداخل و(٢) محتويات كل مدخل (نوع المعلومات، مثل المعلومات الصرفية والنحوية والشرح وأسلوبه ولغته والشواهد والترادف والتضاد والتعابير المركبة التي تدخل فيها الكلمة) إضافة إلى عنصر مهم آخر، وهو(٣) عدد المداخل. فعدد مداخل المعجم يتراوح بين ألفين وخمسة آلاف للمتعلمين من غير العرب ويصل إلى عشرات الآلاف في المعاجم العامة الموجهة لجمهور العرب. كذلك يختلف العدد بحسب موضوع المعجم الخاص، مثل ألفاظ القرآن الكريم. ثم هناك (٤) المنهج المتبع في ترتيب مداخل المعجم (حسب الجذر أو الجذع/ النطق) وهكذا.

٢-٢- تحديد مداخل المعجم ومحتويات المدخل**:**

كان السائد هو اجتهاد واضع المعجم في تحديد المداخل (غالبًا مستعينًا بالجهود السابقة). وهو أمر مقبول في غياب معايير موضوعية لذلك. ولكن المتبع حاليًا في المعاجم الحديثة المعاصرة، في عصر الحوسبة، اعتمادها على المدونات اللغوية التي تجمع آلاف النصوص وملايين الكلمات، وتحليلها إحصائيًا للتعرف على الألفاظ الشائعة والمعاني الشائعة، بدلًا من الحدس والتخمين في اختيار الألفاظ المناسبة.

٢-٣- إعداد المعجم ونشره لتقديمه لجمهور المستفيدين.

في كل من المرحلتين الثانية والثالثة دور للحاسوب يعين المؤلف في إعداد المعجم وفق أسس علمية منهجية. ونوجز في الصفحات التالية أهم استخدامات الحاسوب في كل ذلك.

٣- أهم استخدامات الحاسوب في العمل المعجمي.

٣-١- المدونات اللغوية:

يستفاد من المدونات اللغوية في عدة صور، منها:

٣-١-١- الدراسة المعجمية وصناعة المعاجم lexicography:

لعل الدراسات المعجمية وصناعة المعاجم كانت من أوائل التطبيقات العلمية والعملية للمدونات اللغوية ، حتى من قبل التدوين الإلكتروني أو الرقمي ، كما هو واضح من أعمال مايكل وست Michael West وثورندايك ولورج Thonrndike and Lorge. ونجد شرحا جيدا لدور المدونات المحوسبة في كتابات سنكلير Sinclair الذي كان الباحث الرئيس في مشروع مدونة COBUILD والمعاجم الصادرتحت هذا المسمى، والذي تم بالتعاون بين جامعة بيرمنغهام Birmingham وناشر المعاجم كولينز Collins. (انظرمثلا: *Corpus, Concordance and Collocation* (١٩٩١) .Sinclair, J).

٣-١-٢- دراسات المصطلحية (علم المصطلح terminology):

تعتبر المدونات المتخصصة في مجالات علمية معينة مصدرًا مهمًا لاستخلاص المصطلحات العلمية منها، مثل معرفة المصطلحات المتكررة ومعانيها في سياقات النصوص العلمية المتخصصة. (انظر الدراسة الممتازة التي قدمها عبدالمحسن الثبيتي: **الندوة الدولية الأولى عن الحاسب واللغة العربية : الأوراق البحثية** وكذلك الدراسة المقدمة من عزالدين غازي في الندوة ذاتها، وكتاب **علم المصطلح: مبادئ وتقنيات** Bowker, Lynne and Pearson, Jennifer (٢٠٠٢). *Working with* -ترجمة ريما بركة، وكتاب *Specialized language: A practical guide to using corpora*.(

٣-٢- تحديد الألفاظ الشائعة:

لاشك أن هذا من أبرز أمثلة الاستفادة من المدونات اللغوية المطبوعة والمحوسبة. فقد كانت جميع قوائم شيوع الألفاظ في اللغات المختلفة مبنية على مدونات مختلفة . من هذه القوائم الرائدة في الإنجليزية قوائم مايكل وست Michael West وثورندايك ولورج Thorndike and Lorge . وفي العربية قوائم بريل ولانداو وفاخر عاقل وداود عبده . فقد قام هؤلاء بالاحصاء اليدوي للكلمات في نصوص مطبوعة للتعرف على الكلمات الشائعة فيها . كذلك قام بعضهم بأبعد من ذلك ، حيث قام مايكل وست بإحصاء تكرار المعاني المختلفة للألفاظ semantic count ، كما هو واضح من عنوان كتابه. وفي الآونة الأخيرة نجد عدة أساليب للاستفادة من المدونات اللغوية (المحوسبة حاليًا).

ولكن قبل الخوض في هذا الموضوع لا بد لنا من إيضاح عدة أمورمهمة في هذا المجال، مثل تحديد مفهوم الكلمة والمصطلحات ذات العلاقة، حيث يحدث خلط كبيرلدى بعض الباحثين فيه.

٣-٣- تحديد مفهوم الكلمة والمصطلحات ذات العلاقة**:**

١- الكلمة الحاسوبية : مجموعة من الحروف أو المحارف، أوحتى رمز واحد، مسبوقة بفراغ ومتبوعة بفراغ.

٢- الكلمة الفعلية *token*: الكلمة حيثما وردت وبأيه صورة كانت (مثلا الفعل بصورة ما أو بتصريفاته المختلفة – كتب ، كتبنا ، كتبا ، نكتب ، يكتبون ...أو الاسم في صيغة المفرد والمثني والجمع) . وتحسب تبعًا لعدد مرات ورودها.

٣- الكلمة النوعية type : ويقصد بها الكلمة في صورة ما ، دون النظرإلى تكرارها. لذلك نجد أن عدد الكلمات النوعية أقل من عدد الكلمات الفعلية في النص.

٤- الكلمة بمعنى الصورة الأساس أو الجذع أو الساق stem . (مثلا: الاسم المفرد – كاتب - أو الفعل في صورة الماضي للمفرد المذكر الغائب – كتب.)

ويجدر بالذكر هنا أنه في اللغة العربية تحسب جميع تصريفات الفعل للشخص (أي الفاعل – مثل نكتب، يكتبون، تكتبين...) والزمن (ماض – مثل كتبوا، مضارع- مثل يكتبون، أمر- اكتبا) كلمة واحدة: **كَتَبَ،** وذلك بغض النظر عن إعراب الفعل، مرفوعًا كان أم منصوبًا أم مجزومًا.

كذلك الاسم المثنى (كتابان) والجمع السالم (تلميذات و مسلمون) تحول إلى المفرد، بغض النظر عن الحالة الإعرابية من رفع أو نصب أوجر. فتصبح في حكم الاسم المفرد (**كتاب،وتلميذة ومسلم)** أما جمع التكسير(مثل جبال وأعوام) فيعتبره البعض كلمة مختلفة، حيث إنه ليس قياسيًا.

لبيان ماسبق، يُذكرأن عدد كلمات القرآن الكريم ٥٢٧٧ كلمة مختلفة اي نوعية، بينما مجموع الكلمات الفعلية ١٥٧٩٣٥، وفق إحدى الدراسات الحاسوبية للقرآن الكريم (Qur'an statistics)

٥- الوحدة المعجمية lexeme : قد تكون كلمة واحدة أو عبارة تتكون أكثر من كلمة ، كما في التعبيرات الاصطلاحية (مثلا: ضرب أخماسا في أسداس، وانتقل إلى إلرفيق الأعلى).

٦- الجذع stem (lemma): وهو أبسط صورة للكلمة ، أي دون سوابق أو لواحق، كما ذكرنا في (٤) أعلاه.

٧- الجذر root : في العربية هو الأصل الذي تشتق منه الكلمات ، مثل "غ ف ر" و"ك ت ب".

ويجب أن نفرق بين "الجذع" و "الجذر" في اللغة العربية، فالكلمات "استغفر" ، "استغفار" و "مستغفر" كل منها جذع مستقل ، ولكنها مشتقة من جذر واحد ، هو "غ ف ر".

ويطلق الجذرعلى المكون الأساس للكلمة في اللغات الأوربية (من أمثلة ذلك الجذور اليونانية واللاتينية ، مثل bio, geo, logy).

٨- تجريد الكلمة من الزوائد lemmatization: من أمثلة ذلك إزالة الباءالسابقة والهاء اللاحقة من "بكتابه" ، لنصل إلى الجذع "كتاب".

٩- قائمة الكلمات المستثناه stop list:

عند طلب الكشاف السياقي أو قوائم الشيوع في المدونة ، هناك كلمات لانحتاج إلى معرفة شيوعها (مايسمى بالكلمات الوظيفية ، مثل الضمائر وحروف الجر أو حروف العطف مثلا) ، لكثرتها ولأننا نعرف سلفا أنها أكثر الكلمات شيوعا في اللغة. في هذه الحالة يمكننا أن نعطي للبرنامج الإحصائي قائمة بهذه الألفاظ ليتجاهلها الحاسوب.وهذه هي التي تسمى قائمة الكلمات المستثناه stop list .

١٠- شيوع الألفاظ word frequency: للتعرف على مدى شيوع أي تكرار كلمة معينة نقوم بعملية إحصاء الكلمات word count.

١١- شيوع المعاني semantic count: يشير هذا المصطلح إلى إحصاء المعاني المختلفة للكلمة متعددة المعاني polysemous، وينطبق بصورة خاصة على المشتركات اللفظية homonyms. مثلا ، نجد أن "عين" بمعنى أداة البصر أكثر شيوعًا منها بمعنى عين الماء أو عين بمعنى الجاسوس.

٣-٤- الأدوات اللازمة للاستفادة من المدونات اللغوية:

من المعروف أننا لكي نستفيد من المدونة اللغوية نحتاج إلى برمجيات مناسبة للوصول إلى المعلومات المطلوبة، أهمها: (١) المحلل الصرفي، ويسمى أحيانًا المجذِّع stemmer الذي يجرد لنا الكلمات من اللواصق (كبعض حروف الجر والضمائر المتصلة وعلامات التثنية والجمع مثلًا، وإعادة الأفعال إلى أبسط صورها، كما بينا سابقًا) قبل إحصائها عن طريق (٢) المكشاف السياقي concordance الذي يجري لنا العمليات الإحصائية ويرتب لنا الكلمات وفق شيوعها مثلًا أو ألفبائيًا مع بيان شيوع كل كلمة، كما يقدم لنا الكلمات في سياقاتها المختلفة. (انظر الفصل الخاص بالمدونات اللغوية لمزيد من المعلومات.) ومن الأدوات المتاحة للغة العربية برنامج "غواص Ghawwas" العربي وSketch Engine المعروف بتعامله مع العديد من اللغات التي تشمل اللغة العربية (بل ويأتي معه مدونة للغة العربية).

٣-٥- مجالات البحث التي تتعلق بالجانب المعجمي للغة:

من أهم هذه المجالات مايلي:

١- ما مدى شيوع كلمة ما ، بصورة عامة أو في نصوص مختلفة النوع مثلا؟

٢- ما نسبة شيوع كلمة ما في نصين مختلفين؟

يلاحظ أنه عند إجراء هذا النوع من المقارنة أن نأخذ بعين الاعتبار حجم كل نص . لذلك يفرق الخبراء بين الإحصاء أو العد العام / الخام raw ، أي ذكر التكرار في كل نص بغض النظر عن عدد كلماته ، وهو خطأ . فالمفروض أن نلجأ إلى حساب التكرار النسبي normed count ، أي نذكر تكرار الكلمة في كل ألف كلمة مثلا من النصين (أ) و (ب) ، كأن نقول إن كلمة "مسألة" ترد ٥٠ مرة في كل ألف كلمة من النص (أ) وترد ٥٧ مرة في كل ألف كلمة من النص (ب)، عليه نستطيع القول بأن الكلمة أكثر استعمالًا في النص (ب).

٣- ما مدى شيوع معاني كلمة ما (الإحصاء الدلالي semantic count) في نص ما أو في نصوص مختلفة ؟ نجد مثلا أنه في الاستعمال العام ، ترد كلمة "عين" بمعنى البصر أكثر منها بمعنى عين الماء . كذلك نجد أن كلمة "باب" بمعنى المدخل قد يكون أكثر شيوعا منها بمعنى فصل من كتاب مثلا.

٤- هل للكلمة ارتباط بكلمات معينة (المصاحبات اللفظية – سابقة أو لاحقة)؟ ، مثلا: "جماعة من أهل العلم" و"فريق من اللاعبين" و"قطيع من الماشية" .

٥- هل للكلمة ارتباط بلهجة أو نوع من الاستعمال اللغوي genre أو غيرذلك من العوامل غير اللغوية؟ (مثلا: كلمة "تكوين" بمعنى التدريب و"الشغل" بمعنى العمل في المغرب العربي).

٦- ما مدى شيوع المصاحبات اللفظية المعينة ، بصورة عامة أو في لهجات أو أنواع من الاستعمال اللغوي؟

٧- ما هي مجالات استعمال المترادفات (وسيم ، جميل أو قبيح ، قميء أو كبير، ضخم) وتوزيعاتها ، حسب نوع النص مثلا ؟

٨- الاختلاف بين استعمالات الكلمة في حالتي الإفراد والجمع (مثلا كلمة "حبل" و"حبال" في القرآن الكريم، فهي مجازية في الأولى وحقيقية في الثانية).

٩- التعرف على التعبيرات الاصطلاحية (المسكوكات) ومدى شيوعها عامة أو في أنواع مختلفة من النصوص. ويشمل ذلك الأفعال المتعدية بحروف . (انظرمثلا دراسة وفاء كامل (٢٠٠٧) ومحمد الحناش "برنامج لساني-حاسوبي للتعرف الآلي على التعابير المسكوكة في اللغة العربية" في : مجلة **التواصل اللساني** ، ملحق سلسلة الندوات ، المجلد ٣، سنة ١٩٩٦، ص ٨٩.

١٠- دراسة المصطلحات : التعرف عليها ومعرفة شيوعها في النصوص المختلفة. (انظر Bowker, Lynne and Pearson, Jennifer (٢٠٠٢) ، خاصة الفصلين ٨ و ٩ وانظر كذلك الثبيتي (٢٠٠٧)).

١١- دراسة الجذوروالحروف العربية من حيث تكرارها وشيوعها. ويعتبرعملا علي حلمي موسى (بالتعاون مع إبراهيم أنيس وعبدالصبور شاهين) الخاصين بمعجمي اللسان العربي والصحاح بحق رائدا في هذا النوع من الدراسات (انظر قائمة المراجع) . ومن الأعمال الرائدة في العالم العربي أيضا (المعجم الحاسوبي: إحصاء الأفعال العربية في المعجم العربي) من إعداد محمد المراياتي ويحي ميرعلم ومحمد حسن طيان . ولكن يلاحظ أن هذه الأعمال جعلت بعض المعاجم العربية بوصفها مدونات في إنجاز الدراسات المشار إليها.

ويلاحظ هنا أن إعداد المعاجم (أحادية اللغة وثنائيته)، كما أسلفنا أعلاه ونوضح بالتفصيل أدناه، تستخدم المدونات لأغراض مختلفة ، مثل اختيار ألفاظ المعجم (المداخل) وكذلك حصرمعاني الألفاظ (بناء على الكشاف السياقي والنصوص الأصلية- الشواهد) وفي التمثيل لاستعمالات المداخل المختلفة (من واقع المدونة).

كذلك يستخدم الحاسوب أحيانا في التحكم في لغة التعريف أو الشرح ، كما فعلت لونغمان في معجمها *Longman Dictionary of Contemporary English* ، حيث حددت هذه اللغة بالألفي كلمة الأكثر شيوعا في اللغة.

٤- صناعة المعاجم العربية:

تمثل صناعة المعاجم، كما أسلفنا أعلاه، مثالًا جيدًا لأوجه الاستفادة من الحاسوب في خدمة اللغة العربية. ويشمل ذلك مراحل العمل المختلفة في إعداد المعاجم العربية:

٤-١ اختيار مداخل المعجم ومكوناتها:

تساعدنا المدونات اللغوية في اختيار هذه المداخل ومكوناتها وفق أسس علمية منهجية، وذلك بتوفير النصوص التي نختار منها مداخل المعجم والمعلومات ألأخرى بصورة واقعية. كذلك تفيدنا بعض البرمجيات المرتبطة بالمدونات مثل المكشاف السياقي concrdancer في إجراء تحليل إحصائي للألفاظ والتعابير الكامنة في المدونة وترتيبها بحسب تكرارها، لنختار الأشيع منها مثلًا في إعداد المعجم، إضافة إلى توفير السياقات التي يمكن أن نستخلص معاني الألفاظ منها وكذلك الأمثلة الواقعية لاستخدام تلك الألفاظ، وكذلك المصاحبات اللفظية وما شابه.

جدير بالذكر المعاجم الحديثة التي اعتمدت على المدونات اللغوية، مثل (**معجم العربية المعاصرة، ومعجم التعابير الاصطلاحية، والمعجم العربي للطلاب**).

٤-٢ - إعداد المعجم:

يساعدنا الحاسوب في الإعداد الأولي للمعجم، من حيث مداخل ومكونات كل مدخل (معلومات صرفية كالجذر، نحوية، إملائية، تعريفات، أمثلة للاستعمال، مصاحبات لفظية، تعابير اصطلاحية، مرادفات، أضداد...)، وذلك بالاستفادة من برامج إدارة البيانات data base management حيث يخصص سجل

record لكل كلمة أو مدخل، ويخصص لكل من هذه المكونات حقل field في برنامج إدارة البيانات (قواعد المعلومات).

٤-٣- تقديم المادة المعجمية:

يقدم لنا الحاسوب مزايا عديدة في تقديم المادة المعجمية لكل من الباحث وواضع المعجم، لعل من أهمها:

(١) مرونة التقديم والبحث:

يتميز الحاسوب في المرونة في تقديم المعلومات المعجمية وفي الوصول إليها، فهو يتيح لنا البحث عن الكلمة في صورها وعلاقاتها المختلفة (الجذر، الجذع، المرادف وغير ذلك) بصورة تفاعلية تعتمد على رغبة الباحث، مع إيراد المعلومات المتعلقة بها، وذلك دون حاجة إلى ترتيب المداخل، كما هو الحال في المعاجم الورقية.

(٢)الإحالة:

كذلك يتيح لنا الحاسوب الإحالة إلى المداخل المختلفة بسهولة ويسر، عن طريق مايسمى "النص الخفي hypertext" ، حيث إن كل ما على الباحث فعله للوصول إلى الإحالة المناسبة الضغط على الكلمة ذات العلاقة ليجد المادة المحال إليها أمامه. (في النسخة الورقية، يجد الباحث عبارة مثل "انظركذا"، ثم يذهب يدويًا ليصل إلى المدخل المحال إليه.)

(٣)ترتيب المداخل:

في إعداد النسخة الورقية، يساعدنا الحاسوب في ترتيب المداخل ألفبائيا في دقائق أو ثوان معدودة، بدلًا من الجهد الشاق الذي نحتاج إلى بذله لعمل ذلك يدويًا. ثم يقوم المؤلف بتحويل المادة إلى صورة قابلة للنشرالورقي.

(٤)تحديث البيانات:

يجب أن لا تنسى أن من مميزات تخزين المادة المعجمية حاسوبيًا سهولة التحديث بإضافة مداخل جديدة أوبإجراء التعديلات على الموجود منها.

جدير بالذكر أن بعض دور نشر المعاجم لديها ما يسمى "مدونة المتابعة monitor corpus ، تستفيد منها في متابعة المستجدات في اللغة واستخلاص الألفاظ الجديدة أو الاستعمالات الجديدة ومعانيها.

ولعل من أمثلة الاستفادة من الحاسوب في صناعة المعجم العربي "المعجم العربي التفاعلي" في نسختها الثالثة. وهو منتج عربي شارك فيه عدد من المنظمات العربية (المنظمة العربية للتربية والثقافة والعلوم ومدينة الملك عبدالعزيز للعلوم والتقنية ومؤسسات علمية أخرى في المغرب وسوريا.)

في هذا الصدد يجدر بنا ذكر بنوك المصطلحات العربية التي يمكن اعتبارها معاجم متطورة للمصطلحات، مثل "باسم (البنك الآلي السعودي للمصطلحات) وبنك "مُعربي" التابع لمعهد الدراسات والبحوث للتعريب التابع بجامعة محمد الخامس، وبنك المصطلحات التابع لمكتب تنسيق التعريب في الوطن العربي في الرباط، وبنك أراب تيرم Arabterm الذي أنشأه المكتب المذكور بالتعاون مع جهات أخرى مولتها الحكومة الألمانية.
(انظر: بنوك المصطلحات الآلية.)

المراجع

بركة، ريما (ترجمة، ٢٠١٢) **علم المصطلح: مبادئ وتقنيات** لماري كلود لوم. بيروت: المؤسسة العربية للترجمة.

الثبيتي، عبدالمحسن عبيد (٢٠٠٧) "استخدام ذخائر النصوص لاستخلاص المصطلحات المتخصصة" في : **الندوة الدولية الأولى عن الحاسب واللغة العربية : الأوراق البحثية** ، صص ٣١-٣٨.

حاج صالح، عبدالرحمن (١٩٩٩) "ورقة حول مشروع الذخيرة اللغوية " في : مجلة **اللسان العربي** ، العدد ٤٨.

الحناش، محمد (١٩٩٦) "برنامج لساني-حاسوبي للتعرف الآلي على التعابير المسكوكة في اللغة العربية" في : **مجلة التواصل اللساني** ، ملحق سلسلة الندوات ، المجلد ٣، سنة ١٩٩٦، ص ٨٩.

رشوان، محسن والسعيد، المعتزبالله (تحرير، ٢٠١٩أ) . **تطبيقات أساسية في المعالجة الآلية للغة العربية**. الرياض: مركزالملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

السعيد، المعتزبالله (٢٠١٩) "التحليل المعجمي الآلي" في: رشوان، محسن والسعيد، المعتزبالله (تحرير، ٢٠١٩جـ)، صص ٢١٣-٢٤٠.

شهيد الإسلام، محمد (٢٠١٠). "منهج الخليل بن أحمد الفراهيدي في معجمه "كتاب العين" في: *The Dhaka University Arabic Journal*, Volume-١٢, No-١٣, June, ٢٠١٠.

الصيني، محمود إسماعيل صالح. "بنوك المصطلحات الآلية والمعاجم الألكترونية": **السجل العلمي للندوة الثانية لتعريب الحاسوب (٢٧-٣٠ مارس ١٩٩٤م)** ، ٣١١- ٣٣٤

الصيني، محمود إسماعيل (١٩٨٩). المعاجم في الترجمة الآلية" في: **اللسانيات العربية والإعلامية**. تونس :مركز الدراسات الاقتصادية والاجتماعية، ١٩٨٩م ، ١٩٦. -١٨٣

الصيني، محمود إسماعيل(١٤١٢هـ). نحو معجم عربي للتطبيقات الحاسوبية. في : **السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات المنعقدة في مكتبة الملك عبدالعزز بالرياض (٨-١٢ /١١/ ١٤١٢هـ)** ، صص ٥١١-٥٢١.

عاليم، محمد (١٤٢٨هـ) المعجم العربي في ضوء اللسانيات الحاسوبية. في : **الندوة الدولية الأولى عن الحاسب واللغة العربية : الأوراق البحثية** ، صص ١٥٧-١٦٦.

عبد العزيز، محمد حسن ، محمد يونس الحملاوي والمعتز بالله السعيد طه (٢٠٠٨) المعجم الحاسوبي للغة العربية ، بحث مقدم في: **الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية المنعقد بمدينة الملك عبدالعزيز للعلوم والتقنية في الرياض** (أبريل ٢٠٠٨).

عطية، محمد (٢٠١٩أ) "التحليل الصرفي الآلي للمفردات العربية" في: رشوان، محسن والسعيد، المعتزبالله (تحرير، ٢٠١٩جـ)، صص ٩١-١٣٤.

عطية، محمد (٢٠١٩ب) "التحليل الدلالي" في: رشوان، محسن والسعيد، المعتزبالله (تحرير، ٢٠١٩جـ)، صص ١٨٣-٢١٢.

عمر، أحمد مختار (٢٠٠٨) **المكنز الكبير: معجم شامل للمجالات والمترادفات والمتضادات**. الرياض: مؤسسة التراث.

غازي، عز الدين (١٤٢٨هـ) قواعد المعطيات المعرفية للمصطلحية العربية : مشروع مقترح. في : **الندوة الدولية الأولى عن الحاسب واللغة العربية : الأوراق البحثية** ، صص ٢٣-٣٠.

الغامدي، أيمن أحمد (٢٠١٩) "تقنيات الذكاء الاصطناعي والمعالجة الحاسوبية للمتلازمات اللفظية والتراكيب الاصطلاحية" في الفيفي (تحرير، ٢٠١٩) ، صص ١٢٥-١٦٧.

الغامدي، عبدالله شرف و بدرية سليمان الفرهود (١٤٢٨هـ) أداة ويب معتمدة على عملية التحليل الهرمي للحصول على معجم عربي موحد لتقنية المعلومات . في: **الندوة الدولية الأولى عن الحاسب واللغة العربية : الأوراق البحثية** ، صص ٣٩-٥٠.

فايد، وفاء كامل (٢٠٠٧) المتطلبات اللغوية لمعالجة التعابير الاصطلاحية العربية معالجة آلية، في: **الندوة الدولية الأولى عن الحاسب واللغة العربية : الأوراق البحثية** ، صص ١١-٢٢.

القاسمي، علي (٢٠٠٥) **المعجمية العربية بين النظرية والتطبيق**. بيروت: مكتبة لبنان ناشرون.

القاسمي، علي (٢٠٠٨) **علم المصطلح: أسسه النظرية وتطبيقاته العملية**. بيروت: مكتبة لبنان ناشرون.

مراياتي، محمد ، يحي مير علم ، محمد حسن طيان (١٩٩٦) **المعجم الحاسوبي: إحصاء الأفعال العربية في المعجم الحاسوبي**. بيروت: مكتبة لبنان.

موسى، علي حلمي وعبدالصبور شاهين .(١٩٧٣) **دراسة إحصائية لجذور تاج العروس. باستخدام الكمبيوتر** . الكويت: جامعة الكويت.

موسى، علي حلمي (١٩٧١) **دراسة إحصائية لجذور معجم الصحاح باستخدام الكمبيوتر** . الكويت: جامعة الكويت.

موسى، علي حلمي (١٩٧٢). **دراسة إحصائية لجذور لسان العرب باستخدام الكمبيوتر** . الكويت: جامعة الكويت.

موسى، علي حلمي (٢٠٠١أ) "حوسبة التراث العربي" محاضرة ألقيت في مجمع اللغة العربي الأردني في ١٧ نيسان (أبريل) ، ٢٠٠١.

موسى، علي حلمي (٢٠٠١ب) **ألفاظ القرآن الكريم : دراسة علمية تكنولوجية**. القاهرة: الشركة المتحدة للطباعة والنشر والتوزيع.

الندوة الدولية الأولى عن الحاسب واللغة العربية : الأوراق البحثية (٢٠٠٧). الرياض: مدينة الملك عبدالعزيز للعلوم والتقنية.

نصار، حسين (١٩٨٨) المعجم العربي: نشأته وتطوره،ط ٤. القاهرة: دار مصر للطباعة.

المراجع الأجنبية:

Al-Sulaiti, Latifa(2010). Arabic Corpora. At:
http://www.comp.leeds.ac.uk./latifa/arabic_corpora.htm

Atwell, E. (٢٠١٨). Classical and Modern Arabic Corpora: genre and language change.. In R. J. Whitt (Ed.*), Diachronic corpora, genre, and language change*.. John Benjamins.

Bowker, Lynne and Pearson, Jennifer (٢٠٠٢). *Working with Specialized language: A practical guide to using corpora*. London & New York: Routlege.

Francis, N.W. and Kučera, H. (١٩٧٩). *Brown Corpus Manual: Manual of Information to accompany A Standard Corpus of Present Edited American English for use with Digital Computers, Revised and Amplified*. Available at: http://icame.uib./brown/bcm.html#bc٢

Garside, R., G. Leech and G. Sampson (eds., ١٩٨٧). *The Computational Analysis of English: A corpus-based approach*. London: Longman.

Kennedy, G.D. (١٩٩٨) *An Introduction to Corpus Linguistics*. London: Longman.

Kilgarriff, Adam and Grefenstette (٢٠٠٣)"Web as Corpus" in *Computational Linguistics*, Vol. ٢٩. at: www.kilgariff.co.uk./Publications/٢٠٠٣-KilgGrefesnstette-WACIntro.pdf

McEnry, Tony, Xiao, Richard and Tono, Yukio (٢٠٠٦). *Corpus-Based Language Studies: An advanced resource book*. London & New York: Routledge.

Sinclair, J. (ed., ١٩٨٧). *Looking Up: An account of. the COBUILD Project in lexical computing*. London and Glasgow: Collins ELT.s

Sinclair, J. (١٩٩١) *Corpus, Concordance and Collocation*. Oxford: Oxford University Press

Sawalha, M. and Atwell, E. (٢٠٠٨). "Comparative Evaluation of Arabic Language Analysers and Stemmers" in: Coling ٢٠٠٨: Posters and Demonstrations,. ١٠٧-١١٠.

Svensen, Bo. (٢٠٠٩) *A Handbook of Lexicography: The Theory and Pracitice of Dictioary-Making*. Cambridge: Cambridge University Press.

Thorndike, Edward. L. and Lorge, Irving (١٩٧٢). *The Teacher's Word Book of ٣٠,٠٠٠ Word.* New York: Columbia University, Teachers College Press.

West, Michael (١٩٥٣). *A General Service List of English Words: with Semantic Frequencies and a Supplementary word-list for the Writing of Popular Science and Technology*, ١١th Edition. London: Longman, Green.

Related links

Two generally useful URLs to follow are:

http://www.ling.lancs.ac.uk/monkey/ihe/linguistics/contents.htm

http://devoted.to/corpora

The first of these URLs is an on-line accompaniment to the McEnery & Wilson *Corpus Linguistics* textbook. The second URL is a collection of useful URLs covering a wide range of topics of interest to people using, or interested in using corpora.

The possibility of using the Sara programme, released with the British National Corpus, is best explored by visiting http://www.hcu.ox.ac.uk/BNC/.

التقنية في خدمة المصطلح العلمي العربي

المقدمة:

يلاحظ المتابع للنشاط العربي في مجال المصطلحات العلمية، أنه على الرغم التطور العلمي والتقني في مجال الحاسوب والشبكة العنكبوتية، لا زال العمل في حقل المصطلح العلمي العربي وضعًا وتنسيقًا وتعميمًا ونشرًا يعتمد في كثير من الأحيان على العمل اليدوي وشبه اليدوي وعلى الاجتهادات الشخصية سواء على مستوى الأفراد أو اللجان أو المجامع اللغوية، دون وعي بالذخيرة التراثية في مجال المصطلحات والثروة الكبيرة التي أنتجتها وتنتجها المؤسسات اللغوية والعلمية والمهنية في العالم العربي، مما أدى إلى تضييع كثيرمن الجهد وتكرار المصطلحات، بل وتعارضعها مع بعضها البعض أحيانًا.
ونرى أن أي تطور حقيقي في هذا المجال لا بد أن نستفيد فيه من الحاسوب والشابكة (الإنترنت) بصورة جدية تخدم المصطلح العلمي والعاملين على وضعه وتنسيقه ونشره.

وأقترح خمسة مجالات رئيسة يمكننا فيها الاستفادة من الحاسوب والشابكة، تتلخص فيما يلي:

١ -المدونات اللغوية:

للاستفادة من التراث العلمي الثري في التاريخ الإسلامي والعربي، نحن في حاجة إلى بناء المدونات اللغوية corpora وقواعد البيانات data bases ، وذلك بجمع التراث العلمي العربي وتخزينه في صورة مدونات لغوية مصنفة وفق مجالات تخصصاتها، مثلًا: الفيزياء والكيمياء والطب والرياضيات والزراعة وغيرها من العلوم البحت والتطبيقية، وكذلك في مجالات العلوم الإنسانية كعلوم اللغة والفقه وغيرها. فمن هذه المدونات نستطيع استخلاص الآلاف من المصطلحات التي تزخر بها كتب التراث العلمي العربي عن طريق تحليلها باستخدام أدوات مثل المكشافات السياقية concordancers والمحللات الصرفية، حيث نجد حصرًا للألفاظ واستعمالاتها في سياقاتها الطبيعية ، ونتعرف على شيوعها في المصادر المختلفة عبر العصور.

٢ - قواعد بيانات المصطلحات التراثية:

يضاف إلى تلكم المدونات اللغوية نرى ضرورة بناء قاعدة بيانات للمعطيات المصطلحية التي تزخر بها المعاجم العربية المتخصصة في المجالات (الفنون) المختلفة، مثل كشاف اصطلاحات الفنون للتهانوي، وكتاب جامع الأدوية والأغذية لابن البيطار، ومفاتيح العلوم للخوارزمي، والتعريفات للجرجاني وغيرها مما تزخر به المكتبة العربية التراثية.

وهناك مصدر تراثي ثالث أرى من المفيد توفيره للعاملين في مجال المصطلحات العلمية، ألا وهي معاجم المعاني أو الموضوعات، مثل **المخصص** لابن سيده **وفقه اللغة** للثعالبي، حيث نجد فيها آلاف

المصطلحات وألفاظ الحضارة المناسبة للموضوعات والمسميات المختلفة التي يغفل عنها كثير من اللسانيين وعلماء المصطلحات.

٣ - قواعد بيانات المصطلحات العلمية الحديثة:

نجد أن هناك عددًا هائلًا من هذه المصطلحات قد تصل إلى مئات الآلاف من المصطلحات في بطون المعاجم العلمية التي تصدر من المنظمات العلمية العربية (الطبية، الهندسية، الصيدلية، الإدارية، التربوية، الزراعية...) والمؤسسات العلمية والبحثية (مثل مؤسسة الكويت للتقدم العلمي ومعهد الدراسات والأبحاث للتعريب في المغرب) في العالم العربي، وفي المعاجم أو النشرات المعجمية أو المجلات العلمية التي تصدرها مجامع اللغة العربية وما في حكمها في الدول العربية المختلفة، مثل المجلس الأعلى للغة العربية في الجزائر. هنا نجد الحاجة إلى توفير كل هذا الإنتاج المصطلحي للقارئ والباحث والمصطلحي والمترجم العربي إما في هيئة قاعدة للمعلومات أو البيانات وهو ما نناقشه أدناه. فلا شك أن مثل هذا العمل سيوفر الكثير من الجهد على العاملين في مجال المصطلحات العلمية والمتعاملين معها، مما يساعد في تجنب تكرار الجهود (من ثم تكرار المصطلحات للمفهوم العلمي الواحد أو استخدام المصطلح الواحد لعدد من المفاهيم، مما نشاهده كثيرًا للأسف). ولا شك أن ذلك سيعين غالبًا في تحقيق التنسيق العملي في مجال المصطلحات العلمية.

٤ -١ لبنوك الآلية للمصطلحات العلمية:

من المعروف أن بنك المصطلحات الآلي هو في الواقع نوع من قاعدة البيانات، ولكنها تخضع لمعايير أكثر صرامة ودقة من قاعدة البيانات.

هناك حاليًا عدد من بنوك المصطلحات في العالم العربي، منها بنك لكسار لدى معهد الدراسات والأبحاث للتعريب في الرباط، وبنك المصطلحات الموحدة في مكتب تنسيق التعريب التابع للألكسو في الرباط، وبنك أرابتيرم للمصطلحات التقنية الذي أنشئ بتعاون بين مكتب تنسيق التعريب ودعم من الحكومة الألمانية، والبنك الآلي السعودي للمصطلحات العلمية (باسم) في المملكة العربية السعودية. ولكنها لا تفي بالمطلوب لأسباب مختلفة، منها الشمول (مثل التركيز على المصطلحات "الموحدة" أو مصطلحات التقنية، ودرجة الموثوقية (من حيث مصدر المصطلح) والإتاحة لعامة المستفيدين. كما تختلف من حيث كمية المعلومات المتعلقة بالمصطلحات وأسلوب تقديمها وتحديثها، إلى غير ذلك من مشكلات إدارية وتقنية.

في ضوء ما سبق، أقترح إنشاء بنوك المصطلحات التالية:

٤-١ - بنوك المصطلحات العلمية المحلية:

كنت في مناسبة سابقة اقترحت أن تنشئ كل هيئة معنية بالمصطلحات العلمية (مثل مجامع اللغة العربية) بنكًا آليًّا للمصطلحات التي تضعها أو وضعتها خلال السنوات الماضية، وذلك بمواصفات معينة تساعد على توثيق المصطلحات وواضعيها وعلى مراجعة ما تم إنجازه لتجنب التكرار والتناقض، ولتحقيق التحديث المستمر...إلخ. (انظر" بنوك المصطلحات الآلية: ورقة عمل مقدمة إلى مؤتمر الدورة الخامسة والثمانين مجمع اللغة العربية بالقاهرة" (٢٠١٩/٣/٢٥ إلى ٢٠١٩/٤/٨) وملحقات الورقة المذكورة).

كما أنني تقدمت بورقة عمل خاصة إلى مجمع اللغة العربية بالقاهرة تقدم تصورًا لتنفيذ مثل هذا العمل (ويسعدني توفير نسخة من الدراستين للراغبين فيهما.)

٤-٢- بنك المصطلحات العلمية المجمعي:

في مؤتمر اتحاد مجامع اللغة العربية في القاهرة في العام ٢٠٢٠، تقدمت باقتراح لإنشاء بنك للمصطلحات لتوثيق جميع ما أنتجته وتنتجه مجامع اللغة العربية وما في حكمها في العالم العربي. وقد وافق المؤتمر على المشروع. وكُونت لجنة للعمل عليه. لكن ظروف جائحة الكورونا وانشغال الاتحاد بمشروع معجم اللغة العربية التاريخي حال دون الشروع العملي في مشروع بنك المصطلحات حتى يومنا هذا. وكلنا أمل ألا يتأخر البدء في هذا العمل وإنجازه.

٤-٣- البنك العربي المركزي للمصطلحات العلمية:

لا شك أن هذا يمثل المرحلة الأهم في مجال المصطلح العلمي العربي. وقد دعا إليه عدد من الباحثين من قبل (مثل على القاسمي ومجمع اللغة العربية الفلسطيني، كما دعا إليها مكتب تنسيق التعريب في دراسة مستفيضة بعنوان: "في الحاجة لبنك عربي مفتوح وموحَّد للمصطلحات" صدرت بتاريخ مارس ٢٠٢٠). وهو لا يتعارض مع بنوك المصطلحات الحالية والمقترحة، لأنه من المفترض أن يستفيد منها جميعًا في سبيل الوصول إلى إنشاء بنك عربي متطور بمواصفات أهمها:

- خضوع الجمع فيه على معايير علمية دقيقة تعالج مشكلات التكرار والتعارض وأوجه الاختلاف الأخرى في مجال المصطلحات العلمية العربية. وتشكل لهذا الغرض هيئة أو لجنة علمية متفرغة.

- شمول البنك على المصطلحات في جميع المجالات العلمية البحت والتطبيقية والإنسانية والاجتماعية، وفق معايير موحدة (من حيث البيانات كالتعريف والشرح والترادف والإحالات، والاستفادة من وسائل الإيضاح اللازمة كالصور والرسوم التوضيحية وما شابهها).

- التوثيق الدقيق لمصدر كل مصطلح (مع بيان تاريخ وضعه كلما أمكن ذلك).

- إتاحة الوصول إلى البنك لجميع الراغبين في ذلك من علماء ومؤلفين ومترجمين وغيرهم، وذلك عن طريق ليس الحواسيب فحسب بل وعن طريق الهواتف الذكية كذلك.

- التعريف الجيد والواسع بالبنك وطرائق الاستفادة منه.

- إتاحة ربط المستفيدين ببنوك المصطلحات العالمية (مثل الأمم المتحدة وغيرها من المؤسسات والمنظمات الدولية التي تتعامل مع المصطلحات العربية (انظر الملحق التابع لبحث بنوك المصطلحات الآلية المشار إليه آنفًا وكذلك الدراسة المفصلة التي أعدها مكتب تنسيق التعريب في مارس ٢٠٢٠.)

انظر ورقة مكتب تنسيق التعريب، ٢٠٢٠ لمزيد من التوصيات والمواصفات، ويشمل ذلك "مراحل تنفيذ المشروع" (ص ٢٨-٢٩) وانظر كذلك دراستنا عن "بنوك المصطلحات الآلية" المقدمة إلى مجمع القاهرة في العام ٢٠١٩، لمقترحات حول المكونات التفصيلية لبنك المصطلحات.

ويُقترح أن ينشأ هذا البنك المركزي لدى مكتب تنسيق التعريب التابع للألكسو، فأحسب أن لديه الخبرة المطلوبة. ولكن ذلك العمل يتطلب تمويلًا سخيًّا لا يقل عن الدعم الذي يتلقاه بنك أرابتيرم الحالي الذي أنشاه المكتب بدعم من الوزارة الفدرالية الألمانية للتعاون والتطوير الاقتصادي، حتى يتسنى للمكتب تعيين الخبراء وتشكيل اللجان المتخصصة وتوفير البرمجيات والعتاد الضروري لإنشاء مثل هذا البنك المركزي.

ولكنني أتساءل عن الحكمة في تسمية البنك المركزي الموحد المقترح "أرابتيرم ARABTERM" الذي يحمله البنك التقني المذكور أعلاه.

تعقيب:

لا شك أن من يسمع بعض هذه الاقتراحات يشعر بتكرار الجهد. ولكن الواقع أثبت لنا أنه على الرغم من احتمال تكرار الجهود، فهو مفيد لأن إنشاء بنك عربي مركزي مشروع حضاري طموح، غير أن تكرار الدعوة إليه يثبت لنا عدم تحققه حتى اليوم لأسباب مختلفة. لذلك رأيت اقتراح بنوك المصطلحات من الفئات المختلفة. فهي أيسر تحقيقًا لارتباطها بجهات محددة يمكنها أن تنجزها بسهولة نسبية.

أما البنك العربي المركزي، فيحتاج إلى جهد أكبر وتكلفة أعلى قد لا تتوافر لدى أي من المؤسسات العربية الحالية.

٥- منصة المصطلح العلمي العربي:

حيث إن العمل الجيد يتطلب مشاركة وتعاونًا بين أطراف وأفراد مختلفي الخلفيات والمشارب، فأرى ضرورة إنشاء منصة أو منتدىً على الشابكة للمصطلح العلمي في العالم العربي، ويكون متاحًا لجميع المعنيين والمهتمين به من أفراد ومؤسسات من أنحاء العالم. ويتم في هذه المنصة مناقشة ومعالجة القضايا المختلفة مما يتعلق بالمصطلحات العلمية: إضافة، ترجمة وضعًا، تنسيقًا، نشرًا وتعميمًا. ويشرف على هذه المنصّة البنك العربي للمصطلحات العلمية المذكور آنفًا. فستكون هذه المنصّة رافدًا مهمًّا لإثراء البنك العربي المركزي، بعد إخضاع ما يرد فيها أو ينتج عنها للمعايير اللازمة.

من اطلاعنا على ورقة مكتب تنسيق التعريب، نجد أنه يؤيد مثل هذا العمل، وذلك تحت عنوان "البنك باعتباره منصة سحابية تشاركية مفتوحة" في الدراسة المشار إليها آنفًا. وفي الورقة تفاصيل مهمة يمكننا الاستفادة منها في إنشاء هذه المنصة الافتراضية.

٦- التدريب الافتراضي:

من الملاحَظ أن معظم العاملين في مجال المصطلحات العلمية وغيرهم (كالمترجمين) يضعون المصطلحات بطرائق عشوائية لا تخضع للأسس العلمية اللازمة سواء من حيث أسس وضع المصطلح أم قواعد اللغة العربية، إضافة إلى جهلهم بالمصادر المعجمية المتوافرة.

من هنا نجد ضرورة تقديم دورات تدريبية أو ورش عمل للمصطلحيين والمترجمين العرب على هذه الأسس العلمية العملية، مثل الاستفادة من التراث اللغوي الثري وطرائق وضع المصطلحات الجديدة، مثل الاشتقاق والنحت والتركيب الإضافي والوصفي والاقتراض والترجمة بأشكالها المختلفة، إضافة إلى مفهوم الحقول الدلالية وشبكات المصطلحات، وأساليب الاستفادة من وسائل التقنية المختلفة.

وقد أصبح تنفيذ مثل هذه الدورات التدريبية وورش العمل أمرًا ميسورًا عن طريق العمَل عن بُعد، حيث يمكننا الاستفادة من الخبراء من مختلف العالم، وإتاحة التدريب لفئات مختلفة من الراغبين فيه دون عناء وتكلفة كبيرة.

الخاتمة:

تحدثنا في الصفحات الماضية عن المجالات المختلفة التي يمكننا فيها الاستفادة من التقنية الحديثة في خدمة المصطلح العلمي العربي، وذكرنا ستة ميادين يمكننا خوضها في سبيل ذلك: المدونات اللغوية التراثية وقواعد بيانات المصطلحات التراثية والحديثة وبنوك المصطلحات الآلية ثم منصّة المصطلح العلمي العربي والتدريب الافتراضي، آملين أن يكون هذا العرض الموجز حافزًا لمزيد من النظر في هذا

الموضوع الحيوي ولتنفيذ بعض مما ورد في الورقة، لعلها تكون خطوات في مجال الاستفادة من التقنية في العمل المصطلحي العربي.

بنوك المصطلحات الآلية

المقدمة:

نظرًا للتطور السريع في انتشار المعلومات الجديدة والتطورات المتسارعة في مجالات العلوم والتقنية المخلفة ولضرورة نقلها للشعوب الناطقة بلغات مختلفة، وجدت الهيئات المختلفة العامة والخاصة (مثل الشركات عابرة القارات) ضرورة ملحة للاستعانة بالحاسوب في إيصال هذه المعلومات بالسرعة اللازمة. فكان اللجوء إلى الترجمة الآلية، ولكن هناك مشكلات في تطبيق ذلك حتى في المجالات التقنية والعلمية. مِن ثَم كان النظر فيما يسمّى بالترجمة بمعاونة الحاسوب (بصور مختلفة). ولكن هذا لم يكن كافيًا أو ممكنًا في بعض الحالات.

في نفس الوقت وجدت الهيئات الدولية والدول ثنائية اللغة أو متعددة اللغات ضرورة الاستمرار في الاستفادة من الترجمة البشرية. ولكن هذا، حتى عند توافره، له محاذيره ومشكلاته. وقد لاحظ المشرفون على مشروعات الترجمة في المؤسسات الكبرى (مثل السوق الأوربية المشتركة التي أصبحت تعرف باسم الاتحاد الأوربي) أن المترجم البشري يواجه مشكلات خاصة في مجال المصطلحات العلمية والتقنية التي تتجدد بصورة مستمرة وسريعة. فمعجم المصطلحات التقليدي مهما كان ثريًا لا يواكب التطور السريع في مجال المصطلحات.

كذلك لاحظ خبراء الترجمة التقنية خاصة والمشرفون عليها الاضطراب الكبير في ترجمة المصطلحات والتعامل معها أثناء الترجمة. فنجد المترجم يترجم المصطلح الواحد بمقابلات متعددة في النص المترجم أحيانًا أو يستخدم مصطلحًا واحدًا لأكثرمن مفهوم واحد. ومن أهم أسباب هذه الصعوبات التي يواجهها المترجم أو تشكو منها النصوص المترجمة التي أشرنا إليها أعلاه غياب المراجع المعجمية المناسبة والمواكبة للتطور السريع في مجالات العلوم والصناعة أو صعوبة الوصول إليها.

وبما أن المصطلحات وترجمتها تمثل عصب الترجمة العلمية والتقنية (حتى إن بعضهم ذكر أن حوالي ٤٠% من مشكلات الترجمة في هذه المجالات هي في حقل المصطلحات) نبعت فكرة بنوك المصطلحات الآلية التي تستفيد من التطور الكبير في مجال المعلوماتية وأدواتها الحاسوبية: الحاسوب والشبكات المحلية، ثم الشابكة أو الشبكة العنكبوتية في وقت لاحق.

في الصفحات التالية سنتحدث عن أهم ميزات بنوك المصطلحات، متبوعًا بمواصفات لبنك عربي للمصطلحات: مكوناته ومحتوياته اللغوية والتوثقية وأساليب الحصول على المعلومات (استرجاع البيانات) منها.

ونرفق بالدراسة ثلاثة ملاحق تشمل :

(أ) تعريفًا ببنكي المصطلحات العربيين المفتوحين للاستخدام العام "باسم" و"أرابتيرم".
(ب) قائمة ببنوك مصطلحات عالمية تشتمل على اللغة العربية.
(ت) تعريف موجز بأشهر بنكين دوليين ومعجم موسوعي محوسَب.

١- مميزات بنوك المصطلحات الآلية:

١-١- أهم الفروق بين بنك المصطحات والمعجم التقليدي:

لا أن الفرق كبير جدًّا بين بنك المصطلحات والمعجم الورقي، ويشمل ذلك:

- بنك المصطلحات محوسب، ومن ثَمّ يتميز بسهولة الوصول إلى المعلومات وسرعته (شبه الفورية).
- بنك المصطحات قابل للتحديث المستمر، بينما المعاجم الورقية تبقى على حالها ربما عقودًا دون تحديث.
- سهولة التحديث في بنك المصطلحات، بينما المعجم الورقي يكلف الكثير من الجهد والمال لتحديثه.
- إمكانية الانتشار الواسع جدًّا لمحتويات بنك المصطلحات، عند ربطه بالشابكة التي تتخطى جميع الحدود الجغرافية، مما يجعل الاستفادة منه متاحًا للجميع، على العكس من المعاجم الورقية التي لا تصل إلا إلى عدد محدود من الناس، مهما كان عدد المطبوع منه وكانت وسائل نشره وتوزيعه.

ولعل صعوبات الانتشاروبطأها من أبرز العقبات التي حالت وتحول دون وصول الكثير من المصطلحات العربية إلى جمهور المستفيدين.

١-٢- أهم ميزات بنك المصطلحات:

١- توفير المصطلحات المتاحة ومقابلاتها بلغات مختلفة.
٢- إمكانية تحديث المعلومات بصورة شبه فورية (أي حال توفر الجديد منها ومقابلاتها).
٣- المرونة الكبيرة في التعامل معها، من حيث الحصول على المعلومات. من ثَمّ توفير الوقت والجهد على مستخدميها من مترجمين ومؤلفين وباحثين.

٤- سهولة الوصول إلى المعلومات المطلوبة بطريقة فورية. أما في حالة عدم توافر الشبكة المحلية أو العالمية، فيمكننا توفير البيانات في صور أقراص مضغوطة (تسمى المدمجة) أو حتى الأقراص الفديوية أو ذاكرات الفلاش/ اليوإس بي التي يستطيع قراءتها الباحث والمترجم على حاسوبهما الشخصي.

٥- من الفوائد الجانبية المهمة الإسهام في انتشار المصطلحات وتوحيدها (بصورة غير مباشرة)، حيث إن توافرها بيسر يشجع المؤلفين والمترجمين (في المجالات الإعلامية والتقنية والعلمية والأكاديمية) على الاستفادة منها. ومع زيادة المستخدمين للمصطلحات يزيد انتشارها و كذلك توحيدها، حيث إن من أهم وسائل التوحيد العملية الاستخدام الكثيف لها من قِبَل المستفيدين منها (المؤلفين والمترجمين مثلًا). ويكفينا دليلًا على ذلك انتشار المصطلحات الموحدة في مجال المعلوماتية مثلًا في الولايات المتحدة الأمريكية وبريطانيا على الرغم من عدم وجود مجامع لغوية بأي منهما.كما لاحظنا في الواقع أن مجرد الاتفاق على بعض المصطلحات في المؤتمرات (مثل مؤتمرات تنسيق التعريب) لا يضمن التزام حتى الدول أو الهيئات المتفقة على تلك المصطلحات "الموحّدة" من الناحية التطبيقية والعملية.

٦- سهولة استخراج المعاجم الورقية المتخصصة للراغبين فيها، وذلك عن طريق النشر المكتبي والتقليدي.

٧- فيما يتعلق بالهيئات المسؤولة عن وضع المصطلحات وتقييسها (مثل مجامع اللغة العربية)، هناك ميزات إضافية، أهمها مايلي:

أ- الوصول الفوري للمصطلحات المتوافرة في بنك المصطلحات والتعرف على المصطلحات الجديدة غير المتوافرة في البنك، من ثم التعامل معها بالعمل على وضع المصطلحات المقابلة باللغة الوطنية (عن طريق لجان من أصحاب التخصص العلمي والمترجمين والمصطلحيين).

ب- في حالة ارتباط بنك المصطلحات ببنوك المصطلحات العربية والعالمية، يمكن للقائمين على البنك المحلي الوصول إلى المصطلحات المتاحة لدى بنوك المصطلحات الأخرى ومقابلاتها للاستفادة منها، إمّا بتبنيها أو بوضع مقابلات أخرى (إن كان المتوافر يعاني من خلل فني أو علمي).

ت- اكتشاف التناقض والتكرار الذي يعاني منه بنك المصطلحات بصورة سهلة وسريعة. فعند الاستفسار عن مصطلح أجنبي ما نجد جميع مايتوافر من مقابلات لها في اللغة الوطنية (مثل المترادفات). وعند إدخال المصطلح باللغة الوطنية قد نجد أنه يقابل عددًا من المصطلحات الأجنبية. وهذا خلل علمي في مجال المصطلحات، حيث يجب أن يكون هناك مصطلح واحد للمفهوم الواحد. وكل ذلك يستدعي المعالجة العلمية والمصطلحية، لتصحيح المشكلات أو تفسيرها.

ث- التوثيق العلمي للمصطلحات وواضعيها ومعتمديها وتاريخ ذلك، إضافة إلى المعلومات الخاصة بمصادرالمصطلحات وتاريخ هذه المصادر.

١-٣- المستفيدون واحتياجاتهم:

يشمل المستفيدون من بنوك المصطلحات فئات كثيرة من أهمها:

١-٤-١- قارئ النصوص العلمية والتقنية:

لا شك أن قارئ النصوص العلمية يحتاج أحيانا إلى فهم المفاهيم العلمية المتخصصة ومصطلحاتها. وهذا يعني حاجته إلى شرح لهذه المصطلحات والمفاهيم التي تعبر عنها. وهو مايتطلب وجود شرح أو تعريف بالمصطلح (يفضل أن يرافق ذلك معينات بصرية كالصور الثابتة والمتحركة والرسوم التوضيحية). أحيانًا قد يحتاج القارئ إلى معرفة المقابل بلغة أجنبية (الإنجليزية أوالفرنسية مثلًا) ليفهم المصطلح باللغة العربية مثلًا. وأظن أن كل من قرأ نصوصًا مترجمة في مجالات علمية جديدة وجد الحاجة أحيانًا إلى معرفة المصطلح الأجنبي ليفهم النص. وهو ما يفسر كتابة بعض المؤلفين والمترجمين في المجالات الجديدة خاصة المصطلح الأجنبي بجوار المصطلح العربي.

(من هنا نجد أن بعض الهيئات المشرفة على الترجمة تشترط وجود مسرد أجنبي – عربي وكذلك مسرد عربي أجنبي للمصطلحات الواردة في الكتاب المترجم.)

١-٤-٢- مترجم النصوص المختلفة :

لا شك أن كل من جرّب ترجمة أي نص من لغة ما إلى أخرى سيحتاج إلى معرفة المقابلات العربية للألفاظ (المصطلحات خاصة)، وإلا اضطر إلى وضع مقابل عربي بصورة عشوائية (غالبًا) أو إلى تعريب المصطلح (كتابة الكلمة الأجنبية بالحرف العربي).

نود التنبيه هنا إلى دور المترجمين في وسائل الإعلام المختلفة في وضع مصطلحات كثيرة (قد لا تكون دقيقة) وانتشار هذه الترجمات على حساب ما قد يصدر من هيئات رسمية في وقت متأخر نسبيًا، مثل "قمر صناعي" و"ذكاء صناعي" و"فاكس" و"إنترنت"، مقابل "ساتل"، و"ذكاء اصطناعي" و "ناسوخ" و"شابكة"...).

يبدو أن هذه الفئة كانت من أهم أسباب إنشاء بنوك المصطلحات في كثير من دول العالم، وخاصة في المنظمات والبلدان ثنائية اللغة أو متعددة اللغات.

١-٤-٣- مؤلف النصوص العلمية الحديثة:

ويحتاج هذا المؤلف إلى معرفة المقابلات العربية الصحيحة للمفاهيم التي يحاول التعبير عنها في مؤلفاته. وهنا تكمن حاجته إلى الترجمات العربية المناسبة.

١-٤-٤- هيئات المصطلحات والتقييس:

لا شك أن هذه الهيئات مثل مجامع اللغة العربية في الوطن العربي في حاجة ماسة إلى مصادر للمصطلحات العلمية والتقنية (بل وألفاظ الحضارة كذلك) المستجدة في اللغات المختلفة لإيجاد مقابلات عربية مناسبة لها. ونشرها على أوسع نطاق وبالسرعة الممكنة. لذلك يحتاج هؤلاء إلى مايلي كحد أدنى:

- التعرف على جميع ما تجمع لديها نتيجة جهود السنوات الماضية (مع بيان تواريخ وضعها أو اعتمادها) في مكان واحد، بطريقة ميسورة وسريعة، وذلك إضافة إلى هو متاح من مصطلحات عربية من وضع الهيئات المحلية والعربية والدولية المختلفة (مثل الأمم المتحدة والمنظمات التابعة لها والاتحادات الدولية، بل وبعض الشركات العالمية الكبرى).
- تنقيح المصطلحات المتوافرة لديها، مثل تعديل بعضها وحذف المكرر منها.
- الإحالات بين المصطلحات، خاصة المركب منها لتجنب التناقض بين بعضها البعض أو إزالته، وهو ما يمكن توفيره في بنك المصطلحات جيد التصميم.
- متابعة المستجدات من مصادر مختلفة (تشمل بنوك مصطلحات أجنبية) تستقصيها اللجان المتخصصة والمتعاونين معها من مظانها، مثل البحوث والدراسات والكتب الحديثة.
- وهنا لا بد من معرفة المصطلحات ومحدداتها (تعريفها أو شروح لها باللغتين الاجنبة والعربية، مع التمثيل لاستعمالاتها أحيانًا) حتى يتسنى للمصطلحيين إيجاد المقابلات العربية الصحيحة وفق أسس علمية.

(يجدر بالذكر أن علماء المصطلحية يرون ضرورة وضع المصطلح في سياقه العلمي واللغوي، مثل ربطه بالمصطلحات المرتبطة به ومجالاتها، بوصفه جزءًا من تعريف المصطلح.)

a. نحوإنشاء بنك عربي للمصطلحات:

في ضوء الاحتياجات المشار إليها أعلاه، يمكننا تقديم بعض المقترحات لإنشاء بنك عربي للمصطلحات، قد يكون مقره اتحاد مجامع اللغة العربية. وأود الإشارة هنا أن هذا لا يتعارض مع وجود بنوك محلية أي قطرية، مع مراعاة التنسيق بين جهودها والتعاون المشترك بينها. وهو أمر ميسور في ظل وسائل الاتصال الحاسوبية عن طريق الشابكة.

٢-١-١- مكونات بنك المصطلحات:

لا تختلف هذه المكونات من الناحية التقنية عن أية قاعدة بيانات/ معلومات. فهي تشتمل على:

أ- المكونات التالية:

١- ملف أو أكثر، بحسب الأغراض المطلوب تحقيقها، كما نشير إلى ذلك لاحقًا.

٢- يشتمل الملف على سجلات records لكل مدخل معجمي.

٣- يتكون السجل (المدخل) من:

٤- حقول fields يحتوي كل منها على سمكونات المحتوى اللغوي والمحتوى الإداري المذكورة أدناه.

ب - برمجيات لتخزين البيانات، وتصنيفها، واسترجاعها بصورة تفاعلية، أي حسب رغبة المستفيدين من البنك.

جـ- روابط links بمواقع بنوك مصطلحات أخرى، ويختلف في هذا عن قواعد البيانات المعتادة.

٢-١-٢- المحتوى اللغوي:

أ- المصطلحات العربية:

١- المصطلح العربي، ومرادفاته إن وجدت.

٢- تصنيف المصطلح وفق نظام معين (ديوي أو يودي سي) أوبيان مجال تخصصه.

٣- مختصر المصطلح (إن وجد).

٤- المصطلح بصورته الكاملة للمختصرات.

٥- تعريف المصطلح.

٦- شرح للمصطلح باللغة العربية، مع صور ورسومات توضيحية للمصطلحات التقنية خاصة.

٧- تمثيل لاستخدام المصطلح في سياق توضيحي إن أمكن.

٨- الإحالة إلى المصطلحات ذات العلاقة الوطيدة موضوعيًا و لغويًا، وخاصة بالنسبة للمصطلحات المركبة.

٩- مصدر المصطلح العربي، مع بيان درجة موثوقية المصدر إن أمكن.

١٠- تاريخ وضع المصطلح أو اعتماده من الهيئة أو الهيئات المختصة.

ب- المصطلحات الأجنبية:

١- لمصطلح الأجنبي أو اختصاره (إن وُجِد).

٢- المصطلح الكامل للمختصر.

٣- مرادفات المصطلح إن وجدت، مع بيان مصادرها إن لزم الأمر.

٤- تعريف المصطلح الأجنبي أومثال لسياق نصي لاستعماله.
٥- تحديد مصدر المصطلح وتاريخه، و درجة موثوقيته، إن أمكن.
٦- الإحالات المصطلحية إن وجدت (مثل المصطلحات ذات العلاقة).
(يجدر بالذكر أن التصنيف المذكور في (أ) ينطبق تلقائيًا على المقابل الأجنبي ولغته).

٢-٢- البيانات الإدارية:

١- اللجنة أو الشخص المسؤول عن مراجعة المعلومات المدخَلَة.
٢- تاريخ تخزين المعلومة.
٣- تاريخ تحديث البيانات، إن حدث ذلك.
٤- المسؤول عن التحديث.

٢-٣- لبحث واسترجاع البيانات :

يجب أن يتيح بنك المصطلحات إمكانية البحث فيه واسترجاع البيانات بصورة تفاعلية، تتيح للباحث قرارنوع البحث ومستوى عمقه، من حيث البيانات المطلوبة، مثلا:

أ- المصطلح في لغة ما ومقابلاته في لغة أو لغات محددة (مثلًا المصطلح العربي ومقابله بالإنجليزية أو الفرنسية أو بالعكس، أي البحث بالمصطلح بلغة أجنبية ما ومقابلاته بالعربية وبلغة أو لغات أخرى).
ب- المصطلح وتعريفه أو شرح معناه.
ت- مختصر المصطلح (إن وجد).
ث- المصطلح بصورته الكاملة للمختصرات.
ج- المصطلح بلغة ما ومرادفاته، إن وجدت.
ح- المصطلح ومصدره.
خ- المصطلحات في تخصص (أو تصنيف) معيّن، مرتبّة ألفبائيًّا مع مقابلاتها بلغة أخرى.

يبدو أن خاصية التفاعلية المشار إليها من ضمن التحسينات المهمة التي أجريت على بنوك المصطلحات القديمة المعروفة، طبعًا بالإضافة الزيادة في المصطلحات واللغات (مثل يوروديكاوتم Eurodicautum التابع للاتحاد الأوربي وبنك تيرميوم الكندي. وقد اصبح الأول IETB أي بنك المصطلحات الأوربي التفاعلي. أما الثاني فيطلق عليه الآن تيرميوم بلس Termium Plus).

(انظر: الملحق بخصوص هذين البنكين.)

٢-٤- الملفات الإضافية في بنك المصطلحات:

لاشك أن أي بنك للمصطلحات يجب أن يتصف بالتحديث المستمر لبياناته. وفي سبيل ذلك تحتاج إدارة البنك إلى ملفات إضافية، لعل من أهمها:

١- الملف الخاص للمصطلحات الجديدة التي تحتاج إلى وضع مقابلات لها وتوثيقها لتستفيد منها لجان المصطلحات الفنية في وضع المصطلحات العربية الجديدة. ويتاح لهذه اللجان التواصل مع بنوك المصطلحات العربية والأجنبية ذات العلاقة عن طريق الدخول المباشر إلى مواقع هذه البنوك (إن كان ذلك متاحًا) أو بالتواصل مع المسؤولين فيها.
وهنا يأتي دور العلماء والباحثين المتعاونين مع بنك المصطلحات (لجان المجامع المتخصصة مثلًا والأساتذة الجامعيين)، ممن يزودونه بما يصل إلى علمهم من مستجدات في مجال المصطلحات، وذلك من قراءاتهم للدراسات والمراجع الحديثة في تخصصاتهم.
يجدر بالذكر أن بعض ناشري المعاجم لديهم مدونات خاصة لهذا الغرض، تسمى مدونة المراقبة monitoring corpus، لمتابعة المستجدات المعجمية في اللغة وذلك بغرض الاستفادة منها في تحديث معاجمها المحوسَبة والورقية.
٢- ملف ملاحظات المستفيدين وإسهاماتهم بالمصطلحات الإضافية ومقابلاتها، وماشابه ذلك. وتخضع هذه لمراجعة إدارة البنك ولجانه الفنية للاستفادة منها وتحديث بيانات البنك في ضوئها.

١-٣- الربط مع بنوك المصطلحات والمعاجم الإلكترونية:

من المهام المفيدة لبنك المصطلحات توفير الروابط مع مواقع بنوك المصطلحات الأخرى العربية والعالمية للتيسير على المستفيدين إلى الوصول المباشر إليها.

للمزيد من المعلومات، نورد في ملحق هذه الدراسة عددًا من الملاحق الفرعية، وتشمل:

أ- بنوك مصطلحات عربية مفتوحة.

ب-بنوك مصطلحات دولية تشتمل على اللغة العربية.

جـ - بنوك مصطلحات عالمية متميزة

المراجع:

حجازي، محمود فهمي " دور المصطلحات الموحدة في تعريب العلوم ونشر المعرفة"، **مجلة اللسان العربي**، مكتب تنسيق التعريب، الرباط، العدد ٤٧، ١٩٩٩.

حفار، زين الدين. " واقع المصطلح العلمي العربي في وسائل الإعلام: الإعلام الجزائري نموذجً

https://ahmedalawy.blogspot.com/٢٠١٦/١٢/blog-post_٤٣.html

الخطيب، أحمد شفيق."حول توحيد المصطلحات العلمية"، **مجلة اللسان العربي**، مكتب تنسيق التعريب، الرباط، العدد٤٤، ١٩٩٧.

الحمزاوي، محمد رشاد."رؤية عربية لتوحيد المصطلح العلمي وتقييسه"، **مجلة مجمع اللغة العربية**، ، القاهرة، العدد ٩٠، ١٩٩٩.

زرزار، نيكول (الجامعة الألمانية الأردنية) في مقالها بالإنجليزية عن دور أرابتيرم Arabterm في تقييس المصطلحات التقنية العربية.

https://repozytorium.ukw.edu.pl/bitstream/handle/item/٤٢١٤/Towards a standardized technical Arabic can arabterm rise to the challenges.pdf?sequence=١

http://sina.birzeit.edu/wp-content/uploads/٢٠١٤/٠٤/١٨.an-overview-of-the-ArabTerm-project.pdf

صالح، محمود إسماعيل " فوضى المصطلحات في الكتابات العلمية العربية الأسباب وحلول مقترحة"، **مجلة دراسات مصطلحية**، فاس، المغرب، العدد ٣، ٢٠٠٣.

صيني، محمود إسماعيل. "بنوك المصطلحات الآلية (بنوك المعطيات المصطلحية)"، **مجلة اللسان العربي**، الرباط، العدد ٤٨، ١٩٩٩.

الفاضل، عبدالرحمن بن عبدالعزيز. "البنك الآلي السعودي للمصطلحات "باسم"". "، **مجلة اللسان العربي**، مكتب تنسيق التعريب، الرباط، العدد ٤٧، ١٩٩٩.

القاسمي، علي "دور المصطلح العلمي العربي الموحد في تعريب التعليم العالي"، **مجلة اللسان العربي**، مكتب تنسيق التعريب، الرباط، العدد٥٥ ،٢٠٠٣ .

القحطاني، سعد بن هادي. **التعريب ونظرية التخطيط اللغوي**. بيروت: مركز دراسات الوحدة العربية، ٢٠٠٢.

مدينة الملك عبدالعزيز للعلوم والتقنية، الإدارة العامة للمعلومات. **باسم: البنك الآلي السعودي للمصطلحات العلمية**، الرياض، ١٤٣٢هـ/ ٢٠١١م.

نشرة إخبارية عن المنظمة العربية للتربية والثقافة والعلوم (ألكسو) عن أرابتيرم Arabterm

http://www.alecso.org/newsite/جميع-المقالات/٩٧--٨-١٣٤٠/٢٠١٦_alecso_news ١٢.htm

الملاحق

أ- بنوك مصطلحات عربية

ب- بنوك مصطلحات دولية تشتمل على اللغة العربية

جـ - بنوك مصطلحات عالمية متميزة

د- ورقة عمل مقدمة إلى مجمع اللغة العربية بالقاهرة

أ- بنوك المصطلحات العربية المفتوحة:

١- بنك المصطلحات التابع لمكتب تنسيق التعريب (الرباط):

يعتبر هذا البنك مستودعًا للمصطلحات العربية "الموحّدة" أي التي تم الاتفاق عليها في مؤتمرات التعريب التي يعقدها مكتب تنسيق التعريب في الوطن العربي خلال العقود الماضية. وهو ثلاثي اللغات: العربية والإنجليزية والفرنسية.

www.arabization.org.ma

يجدر بالذكر أن "دليل مكتب تنسيق التعريب : بنك المصطلحات العربية الموحدة" لا يشير إلا إلى النسخ الورقية للمعاجم الموحدة (صص ٢٢-٣٢)

٢- بنك باسم التابع لمدينة الملك عبدالعزيز للعلوم والتقنية (الرياض):

العنوان:basm.kacst.edu.sa

المكونات الأساسية لباسم:

يشتمل البنك على مكونين رئيسين:

٢-١- المكون اللغوي:

المصطلحات العلمية والتقنية ومقابلاتها في أربع لغات (العربية والإنجليزية والفرنسية والألمانية)، غير أن ذلك لا يشمل جميع اللغات الأربع. فجميع المصطلحات لها مقابلات بالعربية والإنجليزية مع تعريفات لبعضها أو أمثلة سياقية لاستعمالها، ولكن بعضها فقط يشتمل على المقابلات الفرنسية والألمانية.

٢-٢- المكون التوثيقي:

يشمل ذلك المعلومات المتعلقة بتصنيف المصطلح ورقمه و بمصادر المصطلحات بلغات مختلفة ودرجة موثوقية بعض المصطلحات، إضافة إلى المسؤول عن إضافة المصطلح ومقابلاتها وعن مراجِع بيانات الإدخال وماشابه ذلك.

٢-٣- المرحلة التمهيدية لإنشاء باسم:

في سبيل إنشاء البنك السعودي الآلي للمصطلحات العلمية:

١- تم إجراء دراسة نظرية لأدبيات بنوك المصطلحات الآلية في العالم في أوائل الثمانينيات من القرن الماضي، عند إنشاء البنك المذكور.

٢- قام فريق من الباحثين السعوديين بزيارة بنوك المصطلحات المعروفة في أوربا في أوائل الثمانينيات، والتي شملت البنوك التالية: بنك Eurodicautum (يوروديكاوتوم) التابع للاتحاد الأوربي في لكسمبورق، وبنك Normaterm (نورماتيرم) الفرنسي في باريس، وبنك المصطلحات العلمية التابع لوزارة الدفاع الألمانية في بون، وبنك TEAM التابع للشركة الألمانية العملاقة سيمنز Siemens في ميونيخ، إضافة إلى زيارة مركز المعلومات المصطلحية ومركز المعلومات المصطلحية Infoterm (إنفوتيرم) في فينا، وكذلك المنظمة الدولية للمقاييس ISO في جنيف.

٣- قام المشرف على مشروع بنك المصطلحات بزيارة لمجامع اللغة العربية في دمشق والقاهرة وعمان، إضافة إلى معهد الدراسات والأبحاث للتعريب، ومكتب تنسيق التعريب في الرباط، وإلى مؤسسةالكويت للتقدم العلمي، بهدف جمع المعلومات والمراجع المعجمية المتعلقة بالمصطلحات العلمية.

٤- تم التواصل مع الاتحادات والمنظمات العربية والدولية المتخصصة في المجالات العلمية والتقنية، مثل منظمة الصحة العالمية واتحاد الاتصالات الدولي وغيرها، للحصول على منشوراتها المعجمية التي تعنى باللغة العربية.

٥- عقدت مدينة الملك عبدالعزيز للتقنية (المركز الوطني السعودي للعلوم والتقنية، سابقًا) اتفاقيتين مع الناشر الألماني Klett وشركة سيمنز Siemens العملاقة للتعاون معهما في مجال المصطلحات العلمية، حيث تكفلت المؤسستان الألمانيتان بتوفير المصطلحات التقنية باللغتين الألمانية والإنجليزية ، وتكفلت المدينة بتوفير المقابلات العربية لها.

٦- أنشئ قسم في مكتبة المدينة مخصصة للمعاجم العلمية والتقنية، إضافة إلى المراجع المعنية بالترجمة الآلية والعامة.

٧- وُضِعت الخطة التنفيذية لإنشاء باسم تشتمل على تحديد الأهداف الخاصة وتحديد أنواع البيانات اللغوية والتصنيفية والتوثيقية، مع تصميم نماذج إدخال البيانات (يسميها البعض "الجذاذات") اللغوية والتوثيقية اللازمة.

٨- بدأ باسم بتخزين حوالي ٦٠٠ مصطلح باللغات المختلفة وفق النموذج المحدد. وتم تجريب النظام من زوايا التخزين (إدخال البيانات) وتصنيفها واسترجاع المعلومات، وذلك للتعرف على المشكلات التنفيذية ومعالجتها.

٩- بدأ تخزين البيانات، وتوفيرها للاستعمال الداخلي، وكان ذلك على الحاسوب المركزي للمدينة، مع استخدام المطارف terminals ثنائية اللغة.

١٠- في المرحلتين الثانية والثالثة تم تطوير برنامج التشغيل ليتيح التعامل مع باسم عن طريق الحواسيب الشخصية، ثم ربط البنك بالشابكة (الإنترنت) ليتيح الوصول إلى محتوياته عن طريقها في أنحاء العالم. (انظر:مدينة الملك عبدالعزيزللعلوم والتقنية ومحمود صيني وعبدالرحمن الفاضل في قائمة المراجع.)

٢-٤- محتويات (باسم):

نقرأ في النشرة التعريفية ببنك باسم بأن البنك "يحتوى على ٤٧٥٠٠٠ سجلّا مصطلحيا في أكثر من ٢١٣ تخصصًا علميا دقيقًا، ويأتي في مقدمتها الموضوعات التي تختص بها مدينة الملك عبدالعزيز للعلوم والتقنية." ونجد في النشرة ٢٢ حقلًا علميًا منها: "علوم الفضاء والطيران، علوم الطاقة، علوم الحاسب... علوم الحياة... علوم طبية.... علوم الهندسة..."

(أنظر:باسم: البنك الآلي السعودي للمصطلحات العلمية. مدينة الملك عبدالعزيز للعلوم والتقينة، الرياض ٢٠١١. كذلك نجد معلومات تفصيلية في مقالة عبدالرحمن الفاضل (المشرف على البنك في ١٩٩٩)، في قائمة المراجع).

٣- **المعجم التقني التفاعلي للمصطلحات التقنية الموحدة:**

نقرأ ما يلي في كتيب التعريف بمكتب تنسيق التعريب بالرباط:

٤- "المعجم التقني التفاعلي ARABTERM بنك للمصطلحات التقنية أنشئ بالتعاون بين الألكسو والوزارة الاتحادية للتعاون الاقتصادي والتنمية في ألمانيا.،ويتولى تنفيذه مكتب تنسيق التعريب بالرباط والوكالة الألمانية للتعاون (GIZ). في نشرته التعريفية بمكتب تنسيق التعريب، نقرأ:

"يهدف هذا المشروع إلى إعداد قاموس إلكتروني موسوعي للمصطلحات مصنف حسب المجالات التقنية والقطاعات الصناعية المختلفة وبأربع لغات، هي العربية والألمانية والإنكليزية والفرنسية، ومتاح لجميع المستفيدين المحتملين مجانًا على العنوان التالي: www.arabterm.ogr.

"وقد تم تصميم وتطوير موقع المعجم بشكل يجعله سهل المنال وملبيًا لاحتياجات المستفدين باللغات الأربع؛ كما يتيح شكله التفاعلي التعديل المستمر في محتواه وإضافة مصطلحات مستخدمة ممنا سيسهم بلا شك في توحيد المقابلات العربية للمصطلحات العلمية والتقنية في العالم العربي.

"تترتب مواد المعجم التقني حسب المجالات التقنية والقطاعات الصناعية المختلفة...

"بالإضافة إلى اللغة الألمانية والإنجليزية والفرنسية، هذا القاموس التقني مقابلات للمصطلحات باللغة العربية، مشفوعة بتعريفات دقيقة وصور توضيحية...

"وبهذا يفوق العدد الإجمالي للمداخل ٣٣٥٠٠ مدخلًا (أي ما يعادل ١٣٤٠٠٠ مصطلحًا رباعي اللغة)..." (دليل مكتب تنسيق التعريب: بنك المصطلحات العربية الموحّدة، ٢٠١٦: صص ٣٨-٤٠).

جدير بالذكر أنه بدأ تشغيل بوابة المعجم على الشابكة (الإنترنت) في ١١ مايو، ٢٠١٠. "وأطلق مشروع ARABTERM في شهر يونيو ٢٠١٥ بوابة خاصة بالأجهزة المحمولة يمكن الوصول إليها انطلاقًا من معظم الهواتف واللوحات الذكية..." (دليل المكتب: ص ٤٠).

تذكر نيكول زرزار من الجامعة الألمانية الأردنية أن معهد غوته والمركزالثقافي المصري الألماني والجامعة الألمانية الأردنية تسهم كذلك في مشروع أرابتيرم، كما تذكر أن المعجم لايقتصر على المصطلحات العربية الموحدة بل تورد بعض المرادفات العربية المحلية (انظر مقالها بالإنجليزية عن موضوع تقييس المصطلحات التقنية العربية في قائمة المراجع.)

ب- قائمة بأهم بنوك المصطلحات العالمية التي بها مصطلحات باللغة العربية:

UNTerm (http://unterm.un.org.)

FAOterm (http://fao.org./faoterm/en/) FOA term portal

UNOGTerm (http://conf.unog.ch/unogterm/)

UNESCOTERM (http://termweb.unesco.org/)

ESCWA (http://un.org/infrmation/infoconfres.asp)

VINTRAS

UN Interpreters (http://un-interpreters.net/glossaries.html)

IUT (Telecommunications) Terms and Definitions

(http://www.itu.int/ITUR/index....)

Microsoft Language Portal (http://microsoft.com/Language/en-US/Default.aspx)

Reverso (http://dictionnaire.reverso.net)

METEOTERM (http://wmo.multicorpora.net/MultitransWebBase/Account.mvc/DirectAccess)

MINEFITRM (http://www.minefiterm.finances.gouv.fr/)

جـ - نماذج لبنوك ومعاجم متطورة للمصطلحات:

١- أي إيه تي إي IATE:

قد يكون بنك المصطلحات (قاعدة بيانات المصطلحات) هذا أكبر بنك للمصطلحات في العالم، حيث يشتمل على مصطلحات في ٢٦ لغة (لغات دول الإتحاد الأوربي. ويشمل موظفوه خبراء في التخصصات المختلفة ولغويين – مترجمين وعلماء المصطلحات). ويورد أحدث إحصاء لمداخله أن به أكثر من مليون ومائة وأربعة وثمانين مدخلا. أما عدد المصطلحات فيتقارب الثمانية مليون مصطلحًا. وهو بنك تفاعلي يتيح للباحث أسلوب البحث فيه، من حيث المصطلح ومقابلاته وحقول المعرفة التي يرغب في البحث فيها...إلى غيرذلك.

٢- تيرميوم بلس *TERMIUM Plus*®:

يعتبر هذا البنك الجهة الرسمية المسؤولة عن المصطلحات في كندا. ويذكر موقع بنك المصطلحات المذكور أنه "أحد أكبر بنوك المصطلحات واللغة في العالم، حيث يعطيك معلومات عن ملايين المصطلحات في الإنجليزية والفرنسية والأسبانية والبرتغالية. ويمكنك العثور على المصطلحات والمختصرات abbreviations والتعريفات وأمثلة الاستعمال في عدد كبير من الحقول المتخصصة. ويعتبر بنك المصطلحات هذا أداة ضرورية لفهم المختصرات الأوائلية acronyms وللتثبت من مسميات الوظائف الرسمية ومقابلاتها باللغات الأخرى..." وهو في صورته الأخيرة بنك تفاعلي يتيح للباحث أسلوب البحث فيه، من حيث المصطلح ومقابلاته وحقول المعرفة التي يرغب في البحث فيها...إلى غيرذلك.

إضافة إلى ماسبق يقدم البنك المشورة في مجال الكتابة الرسمية كذلك. (انظرموقع البنك.)

٣- بابلنت :

يمثّل بابلنت BabelNet أكبر معجم موسوعي يوفر معلومات معجمية وموسوعية للمصطلحات في ٢٧١ لغة وبه شبكة دلالية وأنطولوجيا للمصطلحات. وقد تم تطويره في قسم علوم الحاسبات بجامعة سابينزا في روما. انظر: Roberto Navigli "A Quick Tour of BabelNet ١,١ وموقع الشبكة:https://www.babelnet.com/en/

البحث اللغوي
(مثال لسانيات المدونات اللغوية)

المدونات اللغوية ولسانيات المدونات

المقدمة:

من الأمور التي لا يجهلها أحد أن الحاسوب هو أداة القرن الحالي دون منازع. فقد دخل تقريبا كل بيت ومؤسسة عامة وخاصة، وأدى استخدامه في مجالات الحياة العلمية والعملية إلى تطورات كبيرة في هذه الحقول المختلفة. ومن هذه المجالات ميدان اللغة التي تميز بها الإنسان عن سائرالحيوانات. ويمكننا الإفادة من الحاسوب في النشاط اللغوي بوجوه متعددة، يمكننا تلخيصها في ثلاث صور رئيسة هي: الاستعمال العام (مثل منسق النصوص والترجمة بمعاونة الحاسوب وقواعد بيانات المصطلحات) والاستعمال الخاص (أداة للبحث اللساني، كما هو الحال مع لسانيات المدونات اللغوية واستعمال برمجيات قواعد البيانات في تحليل الأخطاء اللغوية) والبحث اللساني الحاسوبي المتخصص (مثل تطوير البرمجيات المختلفة للتعامل مع اللغة، مثل برمجيات التعرف على الأصوات وتوليفها والترجمة الآلية). (انظر الفصل الأول في هذا الكتاب.)

غير أن أحدث وأهم مجال لعبه الحاسوب في خدمة البحث اللغوي هو مجال المدونات اللغوية corpora ومفرده (مدونة corpus، وهو موضوع هذا الفصل من الكتاب.

ولعل مما يلفت انتباه الباحث العربي أنه على الرغم من مرور نصف قرن تقريبا على أول مدونة إلكترونية في اللغة الإنجليزية وحوالي ثلاثة عقود على البحث اللساني المبني على المدونات المحوسبة فإننا نلحظ قلة إلمام اللسانيين العرب بها وبإمكاناتها غير المحدودة. من هنا نأمل أن تكون هذه الدراسة بمثابة حافز للعلماء العرب للخوص في غمار هذا المجال وفي الإفادة من المدونات اللغوية العربية في بحوثهم اللغوية.

١- تعريف المدونة اللغوية ولسانيات المدونات:

لعل أبسط تعريف للمدونة اللغوية هو: مجموعة من النصوص اللغوية الشفوية والمكتوبة الموثقة (من حيث المصدر والتاريخ والنوع كحد أدنى). ومن التسميات الشائعة، إلى حد ما، اسم "الذخيرة اللغوية"، والتي أطلقها عليها رائد العمل العربي في مجال المدونات المحوسبة اللساني الجزائري عبدالرحمن حاج صالح، غيرأن تعريفه للذخيرة اللغوية كما ورد في دراسته التي نورد مقتبسا منها يختلف قليلا عن المدونة اللغوية بالتعريف الذي ذكرناه (انظر عبدالرحمن صالح (١٩٩٩). كما أن هناك من استخدم "المتون" (مثل صالح العصيمي) و"المكنز" (عبدالغني أبو العزم).

ولكن يبدو أن "مدونة لغوية" هي الأكثر شيوعًا من بين هذه المصطلحات.

٢- مقارنة بين لسانيات المدونات والتحليل اللغوي التقليدي:

قبل أن نتحدث عن أوجه التشابه والاختلاف هذه، يجب أن نؤكد على ما ينبه إليه كثير من اللسانيين، وهو أن لسانيات المدونات بنوعيها (الدراسات المبنية على المدونات corpus-based studies والدراسات الموجهة بالمدونات corpus-driven studies)، ليست نظرية لسانية بل هي طريقة في البحث اللساني. من ثم ليس هناك تعارض، بل هناك تكامل بين النوعين من البحث اللساني. ومع ذلك يمكننا أن نشير إلى اختلافات في التوجه البحثي في كل من البحث اللساني التقليدي conventional ولسانيات المدونات، لعل من أهمها مايلي:-

٢-١- الحدس والواقع اللغوي:

منذ انتقادات تشومسكي الشديدة للسانيات البنيوية structural linguistics واعتمادها على النصوص أو دراسة الأداء اللغوي performance وتأكيده على وجوب أن يكون البحث اللساني وصفا للكفاية اللغوية competence، اتجه البحث اللساني إلى الاعتماد على حدس intuition الناطق الأصلي باللغة والتأملintrospection. وقد كانت حجة تشومسكي آنذاك هزالة المدونات التي كان يعتمد عليها الباحث اللساني (الحجم الصغير للنصوص)، وكذلك امتلاء الأداء اللغوي بمظاهر النقص اللغوي، مثل التردد والتكرار والحذف أثناء التحدث، مما لايعكس أبنية اللغة الصحيحة. في مقابل هذا الاعتماد الكبير على الحدس، نجد أن لسانيات المدونات تؤكد على وجوب الاعتماد بصورة أساسية على النصوص التي تعكس الأداء اللغوي (المدونة corpus)، مع التنبيه إلى حجم المدونات اللغوية الرقمية (التي قد تصل إلى مئات الملايين من الكلمات، بل وبلايين الكلمات) وإلى تمثيلها الحقيقي للغة بشقيها الشفوي والمكتوب، ولكن دون إغفال لدور الحدس والحس اللغوي خاصة في الجانب الكيفي من التحليل اللغوي (التعميم مثلا وتفسير الظواهر اللغوية).

وجدير بالذكر أن مدونات النصوص المكتوبة لا تعاني مما ذكره تشومسكي من عيوب الأداء اللغوي، حيث إن النصوص غالبا ماتكون منقحة. أما ما أشار إليه تشومسكي من عيوب الأداء فهي ترتبط باللغة المحكية. ويلاحظ أن المدونات اللغوية للنصوص المحكية تحتفظ بمثل تلك "العيوب"، ويستفيد منها الباحث اللساني في دراسة خصائص اللغة المحكية، في مقابل اللغة المكتوبة.

٢-٢- ينظرالتحليل اللغوي التقليدي إلى الوحدات اللغوية وأنواعها وخصائصها وإلى التراكيب: مكوناتها ونوعها وتصنيفها وكيفية تكون بعضها من بعض (الجملة وأنواعها وصفاتها والكلمة وتركيبها وحالاتها...).أما الدرس اللغوي المبنى على المدونات، فينظر إلى:

(١)الاستعمال اللغوي، مثلا لماذا نستخدم تركيبا بدلا من آخر (المبني للمعلوم بدلا من المبني للمجهول) في بعض النصوص.

(٢)التركيز على ألفاظ أو تراكيب معينة ودراسة **سياقات استعمالها** (من حيث المستخدم، نوع الخطاب (أدبي، تقريري، علمي)، ومن حيث الجنس الأدبي وغيرذلك.الارتباط association (أ) اللغوي-اللغوي (مثلا المعجمي-المعجمي، كما في المصاحبات أوالمتلازمات اللفظية، واللفظي-التركيبي، كما في أنواع الأفعال الناقصة و اللازمة والمتعدية بأنواعها وما يرتبط بكل منها من تراكيب نحوية). (ب) اللغوي-غير اللغوي (مثل ارتباط استعمال كلمة أو تركيب معين بمؤلف أو مجال أو لهجة أو عصر تاريخي... وهكذا).(انظر Biber, Conrad and Reppen, ١٩٩٨: ٥-١٢).

لكن هذا لا يعني، كما أسلفنا، أن هناك تناقضا بين الدرس اللغوي التقليدي والمبني على المدونات. فالباحث اللساني لابد أن يكون لديه إلمام جيد بالنظريات اللسانية الحديثة في مجالات الأصوات phonology والتراكيب morphology and syntax والدلالة semantics والمعجمية lexicography، إضافة إلى النظريات الخاصة بتحليل الخطاب discourse analysis وأعمال الكلام speech acts والتداولية pragmatics. فلسانيات المدونات كما أسلفنا أعلاه إنما، هي طريقة في البحث اللساني يستفيد من الإمكانات الهائلة للحاسوب والمدونات اللغوية corpora.

٣- مزايا لسانيات المدونات (المحوسبة):

تتسم لسانيات المدونات المحوسبة بمزايا كثيرة، لعل من أهمها مايلي:

- إنها عملية وتجريبية، مبنية على نصوص حقيقية للاستعمال اللغوي، وليس على الحدس الشخصي. وتدرس نماذج واقعية للغة أو التركيب.
- الحجم الكبير للنصوص أوالمدونات التي تخضع للبحث والتحليل.
- التنوع المبنى على أسس علمية لنصوص المدونة لتمثل استخدامات اللغة المختلفة، وذلك بمراعاة التمثيل الأفقي (الجغرافي) والعمودي (التاريخي) والنوعي (الأسلوبي مثلا) للغة واستعمالاتها المختلفة. ولكن ذلك يعتمد بالطبع على التصميم والاختيار الجيدين من قبل معد(ي) المدونة اللغوية.
- استخدام الحاسوب في التحليل بصورة واسعة، آليا (دون تدخل بشري) وتحاوريا (بين الحاسوب والباحث) حيث يقوم الحاسوب بالجانب الآلي، كالإحصاء والتنبيه إلى ظواهر معينة كاللبس المعجمي أو التركيبي، ويقوم الباحث اللغوي بالجانب الذي يتطلب تدخلا بشريا مثل فك اللبس وتحديد الاسم الذي يعود إليه ضمير ما في النص، مثلا).

- سهولة الحصول على النصوص والمدونات (بصور مختلفة: أقراص مضغوطة CD، على الشابكة (الإنترنت)) أو في صورة نصوص رقمية، وباستخدام قارئة المحارف البصرية OCR (لتحويل النصوص المطبوعة إلى نصوص رقمية)، مما جعل بناء المدونات أمرا ميسورا.

(يجب هنا التنبه إلى وجوب مراعاة حقوق الملكية الفكرية لبعض النصوص.)

- الاعتماد على أساليب التحليل الكمية (الإحصائية، مثل شيوع كلمة أو تركيب ما في نص أو نصوص مختلفة) والكيفية (التفسير والاستنتاج) مثل تفسيرسبب كثرة استعمال كلمة أو تركيب ما في أنواع النصوص المختلفة). (انظر١٩٩٨ ,Biber, Conrad and Reppen)
- إمكانية التثبت من صحة القواعد المبنية على الحدس والتخمين في ضوء الاستعمال الواقعي للغة، من ثم التعرف على دقة تلك القواعد من حيث وصفها للاستعمال اللغوي.
- الكشف عن مزيد من المعلومات المعجمية والنحوية وغيرها في ضوء النصوص المتاحة في المدونة. فقد ثبت مثلا في البحث المبني على المدونات أن كثيرا من الألفاظ لها معان تزيد عما يرد عادة في المعاجم المعروفة. (انظر (٩٩-٩٦ ,*Hunston (2002) Corpora in Applied Linguistics*).
- إمكانية الاستفادة من المدونات في شتى فروع المعرفة اللسانية، مثل اللسانية النفسية والاجتماعية والدراسات المعجمية والنحوية وتحليل الخطاب والتداولية، كما سنبين لاحقا.

٤- أنواع المدونات اللغوية:

يمكننا تصنيف المدونات اللغوية من زوايا مختلفة، كالعموم والخصوص والمعاصر والتاريخي،ومن حيث الوسم و عدمه...

أ- من حيث العموم:

١- المدونة العامة:

هي مدونة أنشئت لأغراض مختلفة ومن نصوص متباينة. ومن أمثلتها: "مدونة اللغة العربية" و"المدونة الوطنية البريطانية British National Corpus ".

٢- المدونة الخاصة بنوع من النصوص (نص ديني، شعر، نثر، قصة، كتابة تقنية أو علمية...) أوخاصة بمنطقة جغرافية (بلد أو منطقة معينة)، أوخاصة بمؤلف ما، أوخاص بنص ما ("مدونة القرآن الكريم Qur'an Corpus").

٣- المدونة المقارنة أو المتشابهة comparable:

هي المدونة التي تشتمل على نصوص متشابهة من حيث المحتوى (مثلا رسائل جامعية أو مقالات علمية في مجال ما) بلغتين أو صورتين مختلفتين للغة نفسها. ومن أمثلة ذلك المدونات التي تهدف دراستها للمقارنة بين كُتّاب مختلفين في كل

من اللغتين أو نوعية اللغة، مثل أساليب أهل اللغة ومتعلّميها من غير أهلها. وهذا النوع من المدونات هو السائد في الدراسات اللغوية التقابلية contrastive.

٤- المدونة المتوازية parallel:

تشتمل على مجموعة من النصوص المتماثلة بلغتين مختلفين (مثلا أحد النصين ترجمة للنص من لغة أخرى). ويستفاد من هذا النوع من المدونات في دراسات الترجمة والمترجمين.

وجدير بالذكر أن مشروعات وبرمجيات الترجمة الآلية المبنية على أسس إحصائية (مثل برمجيات شركة آي بي إم IBM) تعتمد بشكل أساس على هذا النوع من المدونات. كما أن هناك برنامجا على الشابكة لتطوير برمجيات الترجمة الآلية المبنية على المدونات المتوازية: Moses Statistical Machine Translation System (انظر قائمة المراجع.)

٥- مدونة المتعلمين learner (متعلّمي اللغة خاصة من غير أهلها):

تتكون المدونة من أعمال أنتجها متعلمون للغة (تلقائية spontaneous أو مستنطقة elicited). وتفيد هذه المدونات في تحليل أخطاء الدارسين من خلفيات ومستويات مختلفة، من ثم تفيد أيضا في دراسات اكتساب اللغة الثانية. ومن أمثلة هذه المدونات:

.Cambridge Learners' Corpus, Longman Learners' Corpus

وللغة العربية مدونات للغة المتعلمين، إحداها التي أنشأتها الباحثتان الماليزيتان (انظر Hassan, Haslina and Nurahihan Mat Daud, ٢٠١١) والأخرى تسمى "مدونة متعلمي العربية المكتوبة The L٢ Written Arabic Corpus التي تم إنشاؤها في جامعة أريزونا بالولايات المتحدة الأمريكية بمركز المصادر التعليمية عن الثقافة واللغة ومحو الأمية (CERCLL- Center for Educational Research in Culture, Language and Literacy) . وتعتبر هذه المدونة التي تتكون من ٣٠٠ مقالة نموذجا لما يمكن أن يكون عليه هذا النوع من المدونات، حيث يذكر الباحثون ما يلي عنها:

"تم وضع المقالات المطبوعة في قاعدة بيانات قابلة للبحث، حيث تم وسمها بمستوى المتعلم (مبتدىء ومتوسط ومتقدم) ونوع المتعلم (دارس لغة ثانية أو التراث) ونوع النص genre (وصف وسرد وتعليمات)" (انظر:http://l٢arabiccorpus.cercll.arizona.edu). وجدير بالذكر أن الفصل الثاني من الكتاب الذي حرره صالح العصيمي (٢٠١٥) قد خصص للحديث عن مدونات المتعلمين.(انظر: عبدالله الفيفي "مدونات المتعلمين" في الكتاب المذكور.

٦- المدونة التعليمية pedagogic:

وتشمل المواد التعليمية لمنهج ما في مدونة واحدة، لأغراض تعليمية مختلفة. (انظر أيضا O'Keeffe, Anne, McCarthy, Michael and Carter, Ronald (٢٠٠٧)). ويستفاد من هذه المدونات في تعليم اللغة وفي فحص المواد التعليمية ودراستها.

ب- من حيث الزمن:

٧- مدونة تاريخية historical/ diachronic:

وهي المدونة التي تشتمل على نصوص من عصور مختلفة، بصورة علمية متوازنة، للتعرف على التطور الذي طرأ على اللغة وعلى استعمالات ألفاظها ومعانيها وعلى تراكيبها إلى غير ذلك. وهذا النوع من المدونات يجب أن يكون الأساس لأي عمل معجمي تاريخي يعتمد على الاستقراء المنهجي، بدلا من الاجتهادات الشخصية والحدس والتخمين. ومن أحسن الأمثلة لهذا النوع من المدونات: ARCHER و تشتمل على نصوص للإنجليزية البريطانية والأمريكية للفترة مابين ١٦٥٠ حتى ١٩٩٠. يجدر بالذكر ما قام به الباحثون في مشروعي المعجم التاريخي للغة العربية (مشروع الدوحة ومشروع اتحاد المجامع اللغوية العربية) مما يمثل مثالًا جيدًا للمدونة اللغوية التاريخية.

٨- مدونة المتابعة monitor

يقصد بالمتابعة متابعة التطور في لغة ما، المعجم خاصة:

نجد أن بعض مؤسسات النشر المعجمي، مثل داركولينز البريطانية Collins، تهتم بمثل هذا النوع من المدونات التي ترصد المستحدثات المعجمية خاصة، عن طريق التحديث المنهجي للمدونة (قاعدة البيانات اللغوية)، وذلك بغرض تحديث معاجمها بين الفينة والفينة وإضافة ما يطرأ على اللغة من ألفاظ جديدة أو ملاحظة الاستعمالات الطارئة لبعض مداخل تلك المعاجم. (انظر (Susan Hunson, ٢٠٠٢: ١٥-١٦)

جـ- من حيث نوع اللغة:

٩- اللغة المنطوقة/ المحكية، مثل نصوص هاتفية، أو تعليقات إذاعية، ومحادثات يومية. وهناك عدد محدود من هذه للهجة العربية المصرية. ويجب التنبيه إلى الفرق بين المدونة البسيطة التي تحول النص المنطوق إلى نص مكتوب فحسب وبين المدونة التي تسجل بصورة دقيقة النص الشفوي (من حيث الوقفات والنبر والتنغيم، كذلك مواقع الترردد والتكرار... إلى غير ذلك مما يحدث أثناء الكلام). ولكل نوع أغراضه وفوائده. (من أمثلة هذا النوع من المدونات:

مدونة بيرغن لإنجليزية مراهقي لندن (كولت) وكذلك مدونة الإنجليزية المنطوقة للمحترفين.

Bergen Corpus of London Teenage English (COLT), Corpus of Spoken Professional English و CALLHOME Corpus، والأخيرة خاصة بمحادثات هاتفية باللهجة المصرية.

هذا ويفضل أن تتوافر تسجيلات النصوص الصوتية للباحثين للرجوع إليها عند الحاجة.

١٠- مدونة اللغة المكتوبة، وهذا هو الغالب على المدونات اللغوية المختلفة.

ولكن هناك مدونات تجمع بين اللغة المحكية والمكتوبة، مثل "المدونة الوطنية البريطانية BNC".

د- من حيث الوسم:

١١- مدونة خام: ويقصد بها المدونة التي تخلو من أية تحشيات annotations (صرفية أو نحوية أومعجمية مثلا). وينطبق هذا على معظم المدونات اللغوية المعروفة، خاصة العربية منها.

١٢- مدونة موسومة tagged: وهي المدونة التي بها تحشيات (صرفية أو نحوية أومعجمية مثلا)، حيث نجد وسما للكلمات مثلا، يبين قسم الكلام الذي تنتمي إليه POS (part of speech) . ومن خير الأمثلة على هذا النوع من المدونات: Lancaster Parsed Corpus (مدونة لانكاستر المعرَبَة للغة الإنجليزية). في الملحق (أ) نجد أمثلة لممدونات عربية موسومة.

وللباحثة سلوى حمادة تقسيمات أكثر تفصيلا في دراستها: "المدونات النصية ودور اللغة العربية في التعامل معها."

ونورد في الملحق (أ) قائمة بأهم المدونات اللغوية العربية.

٥- مواصفات المدونات اللغوية:

من الأمور التي يجب أن تؤخذ بعين الاعتبار عند تصميم المدونات وتقويمها ما يلي:

١-١- الحجم الكبير

غالبا مايكون ذلك ملايين الكلمات للمدونات العامة. ولكن ذلك قد لاينطبق بالضرورة على المدونات الخاصة.

٤-٢- الشمول وتمثيل استعمالات اللغة representativeness:

يعني ذلك أن تشتمل المدونة على نصوص تمثل استعمالات اللغة المختلفة (الشفوية والمكتوبة)، في ضوء الهدف من إنشاء المدونة. فلا تقتصر المدونة اللغوية على استعمال أو أسلوب أو منطقة جغرافية معينة مثلا.

٤-٣- التوازن:

من شروط المدونة الجيدة أن يكون هناك توازن بين أنواع أو فئات النصوص والتخصصات وغير ذلك مما يشمله معيار التمثيل، فلا يطغى مؤلف أو لهجة أو جنس أدبي... على غيره. (انظر:٥٧-٢٤٣ :٨/٤ ,٢٠٠٣ ,.Biber, D).

ولأمثلة لتطبيق هذه المعايير، انظر مثلا .١٧ :٢٠٠٦ ,McEnery et al ;٢٠٠٣ ,Biber في وصفهم للمدونة الوطنية البريطانية British National Corpus وانظر أيضا مانورده لاحقًا عن المدونة اللغوية العربية التي أعدته مدينة الملك عبدالعزيز للعلوم والتقنية في الرياض، والمدونة اللغوية العربية الدولية التي أنجزت في مكتبة الإسكندرية.

على سبيل المثال، نجد أن مدونة براون الرائدة اشتملت على خمسمائة "نوع/ جنس أدبي genre" مختلف من الكتابات المنقحة edited، يمثل كل منها نصًا يتكون من ٢٠٠ كلمة (المجموع: مليون كلمة من النصوص المنشورة). وفي المدونة الوطنية البريطانية، تم التأكيد على أن "الكلام الذي يجمع يؤخذ من عينة تراعي التوازن من حيث العمر والجنس والطبقة الاجتماعية والمنطقة اللهجية " (Meyer, ٢٠٠٢: ١٨). ونورد أدناه بعض التفاصيل الخاصة لهذه المدونة الرائدة في مجال تصميم المدونات اللغوية:

مثال: مدونة براون Brown Corpus:

أشرف على إعداد هذه المدونة لساني أمريكي فرانسس Francis وحاسوبي Kocera في جامعة براون الأمريكية. وقد نشر الكتاب الخاص بهذه المدونة وبعض نتائج دراستها في عام ١٩٦٧ (للإنجليزية الأمريكية). ونظرا لأهمية هذه المدونة وأسبقيتها (مما جعلها مثالا يحتذى به في تصميم كثيرمن المدونات)، فنورد بعض التفاصيل المهمة عنها.

- الهدف: دراسة اللغة الإنجليزية الأمريكية المعاصرة. ونقرأ في دليل المدونة المنقح والموسع في عام ١٩٧٩بعض التفاصيل المهمة التالية:

"تشتمل هذه المدونة المعيارية للغة الإنجليزية المعاصرة من ١,٠١٤,٣١٢ كلمة من النثر الإنجليزي المنقح المنشور في الولايات المتحدة في العام ١٩٦١. وقد حاولنا التثبت قدر الإمكان من كون المؤلفين من الناطقين الأصليين للإنجليزية الأمريكية....

"تم التوصل إلى الفئات الرئيسة وفروعها في مؤتمر عقد في في جامعة براون في فبراير ١٩٦٣. كذلك قدم المشاركون في المؤتمر آراءهم الخاصة بخصوص عدد العينات في كل فئة، ثم استخرج المتوسط لهذه الأرقام واستخدمت في المدونة. هذا وقد أجريت تغييرات طفيفة في وقت لاحق في ضوء التجربة العملية في اختيار العينات. كذلك تم تقسيم العينات إلى فئات أدق بناء على نسب النشر الحقيقي في العام ١٩٦١."
(*Brown Corpus Manual*, ١٩٧٩: ١)

- المحتوى: اشتملت على حوالي **مليون كلمة** بالشكل التالي:

تم تحديد أنواع النصوص المطلوب اختيار عينات منها والفترة الزمنية، في ضوء مناقشات مؤتمر عقد لهذا الغرض، كما ذكر أعلاه. وقد قسمت العينات إلى ثمانية عشر فئة رئيسة، تندرج تحت كل فئة عدد من الفروع.

- الفترة الزمنية كانت عاما واحدا، هي العام ١٩٦١.
- النصوص: ٥٠٠ نص لتمثيل الأنواع المختلفة المذكورة أدناه. ويشير مؤلفا الدليل إلى أنه تم استبعاد الشعر لخصوصية لغته وكذلك الأعمال المسرحية (لأنها تمثل نوعا من اللغة المحكية)، حيث إن الهدف كان دراسة لغة النثر فحسب.(مقدمة الدليل). تم اختيار ٢٠٠٠ كلمة من كل نص.نورد أدناه ترجمة لفئات النصوص الرئيسة والفرعية الخمسمائة كما وردت في دليل المدونة:

أولا: النشر الإعلامي/ المعلوماتي ٣٧٤ عينة.

أ- تقارير صحفية:

سياسية ١٠ يومية + ٤ أسبوعية

رياضية ٥ يومية + ٢ أسبوعية

اجتماعية ٣ يومية

أخبار مهمة ٧ يومية + ٢ أسبوعية

مالية ٣ يومية + ١ أسبوعية

ثقافية ٥ يومية + ٢ أسبوعية

المجموع ٤٤ عينة

ب- صحافة: التحرير

مؤسساتية ٧ يومية + ٣ أسبوعية

شخصية ٧ يومية + ٣ أسبوعية

خطابات إلى رئيس التحرير ٥ يومية + ٢ أسبوعية

المجموع ٢٧ عينة

جـ - صحافة: نقد (المسرح، الكتب، الموسيقى، الرقص)

١٤ يومية + ٣ أسبوعية

المجموع ١٧

د- دين:

كتب ٧

٦ دوريات

مختارات tracts ٤

المجموع: ١٧

هـ - مهارات وهوايات:

كتب ٢

دوريات ٣٤

المجموع ٣٦ عينة

و- أدب شعبي

كتب ٢٣

دوريات ٢٥

المجموع ٤٨ عينة

ز- أدب راق وسير ذاتية ومذكرات... إلخ.

كتب ٣٨

دوريات ٣٧

المجموع ٧٥ عينة.

ح – منوعات:

وثائق حكومية ٢٤

تقارير مؤسسات ٢

تقارير صناعية ٢

دليل جامعي ١

صوت مؤسسة صناعية ١

المجموع ٣٠ عينة.

ط – علوم:

علوم طبيعية ١٢

طب ٥

رياضيات ٤

علوم اجتماعية وسلوكية ١٤

علوم سياسية وقانون وتربية ١٥

إنسانيات ١٨

تقنية وهندسة ١٢

المجموع ٨٠ عينة.

ثانيا: نثر خيالي: ١٢٦ عينة.

ي – رواية عامة:

روايات ٢٠

قصص قصيرة ٩

المجموع ٢٩ عينة.

ك- قصص بوليسية:

روايات ٢٠

قصص قصيرة ٤

المجموع ٢٤ عينة

ل- الخيال العلمي

روايات ٣

قصص قصيرة ٣

المجموع ٦

م- مغامرات وقصص الغرب الأمريكي (الكاوبوي):

روايات ١٥

قصص قصيرة ١٤

المجموع ٢٩ عينة

ن- قصص الرومانسية والحب:

رواية ١٤

قصص قصيرة ١٥

المجموع ١٩

س- الفكاهة:

رويات ٣

مقالات وماشابه ٦

المجموع ٩

العدد الإجمالي للعينات: ٥٠٠ عينة

جدير بالذكر أن المدونة المذكورة متاحة في خمس نسخ، بنفس المحتوى اللغوي ولكن مع اختلافات في بعض الجوانب الإضافية لعل من أهمها ماأضيف إلى النسخة C من وسم نحوي للكلمات (معلومات تتعلق بأقسام الكلام الرئيسة والفرعية). انظر دليل المدونة. (القسم "٤: النسخة الموسومةThe Tagged Version ." من دليل مدونة براون.)

٤-٤- معايير إضافية

من الدراسات المهمة التي عالجت معايير تصميم المدونات اللغوية مقالة سو أتكنز وزميليها Sue Atkins et al.(١٩٩٣) بعنوان "معايير تصميم المدونة اللغوية Corpus Design Criteria " والذي نشرته مجلة الحوسبة الأدبية *Literary Computing* واللغوية (انظر: Atkins, Sue et al., ١٩٩٣). وكذلك مجموعة الدراسات التي وردت في الكتاب الذي حرره مارتن وين Wynne بعنوان: تطوير المدونات اللغوية: دليل للممارسات الجيدة (*Developing Linguistic Corpora: a Guide to Good Practice*) والذي يجمع بين دفتيه مقالات مهمة لعدد من رواد العمل في هذا المجال، وقد صدر في العام ٢٠٠٥.

٦-إعداد المدونة و طرق الاستفادة منها:

هذا وقد يتم جمع النصوص للمدونة قد يتم بطريقة عشوائية أو منتظمة (وفق أسس محددة). كما يتم الاستفادة منها كذلك بصور مختلفة، كما فعل مثلا اللغويون العرب الأوائل عند وضع قواعد اللغة العربية ومعاجمها، ودليل ذلك الاستشهادات التي نجدها في بطون كتب النحو والمعاجم. (انظر أحمد مختار عمر "مصادر اللغويين العرب" في كتابه الموسوم (**البحث اللغوي عند العرب**...) كذلك نجد مثالا على ذلك ما فعله الباحثون في علم اللغة التاريخي في القرن التاسع عشر، و ما فعله اللسانيون البنيويون من أمثال بلومفيد وسابير في دراساتهم للغات الهنود الحمر وغيرهم ممن عملوا في اللسانيات الميدانية في النصف الأول من القرن العشرين، وأخيرا

ونجد أن الجمع بين المدونات والاستفادة منها يتم في صور مختلفة، منها:

- مدونات مطبوعة وكان البحث فيها يدويا:

ما فعله مايكل وست Michael West ومن قبله ثورندايك ولورج Thorndike and Lorge.

من الأمثلة العربية أعمال لانداو، بريل، فاخر عاقل، داود عبده (دراسة في... وقائمة بثلاثة آلاف أشيع كلمة في اللغة العربية) (انظر: عبده، داود، ١٣٩٩).

- المدونة محوسبة ولكن البحث يدوي:

والمثال على ذلك (دراسات كرسي الجزيرة في جامعة الأميرة نورة المذكورة في قائمة المراجع.)

- مدونة محوسبة وبحث حاسوبي محض أو حاسوبي تفاعلي (تعاون بين الحاسوب والباحث):

من أمثلة ذلك مدونة العربية ArabiCorpus التي أعدت في جامعة بريغهام يونغ BYU أشرف عليها المستشرق الأمريكي ديل باركنسون Dell Parkinson في ولاية يوتا بالولايات المتحدة الأمريكية، والتي يصحبها برنامج للتحليل الآلي للمدونة. ومن هذه المدونة استخلِص معجم الخمسة آلاف كلمة الشائعة الذي أعده المذكور مع الباحث بكوولترBuckwalter (انظر قائمة المراجع).

لكن نود أن ننبه إلى أن "المدونة" في سياق لسانيات المدونات بقصد بها المدونة اللغوية المحوسبة، أي المخزنة رقميا في الحاسوب. لذلك نجد بعض الذين يتحدثون عن لسانيات المدونات باسم لسانيات المدونات الإلكترونية electronic corpus linguistics.

وجدير بالذكر في هذا المجال أن الشبكة العنكبوتية (الإنترنت) أو الشابكة كما يسميها البعض يمكن اعتبارها مدونة لغوية شاملة، كما يشيرالباحثان كيلغاريف وغرينفنستيت Adam Kigarriff and Greogroy Grefenstette في بحثهما المعنون الشبكة بوصفها مدونة Web as Corpus حيث يقولان في مقدمة بحثهما أن الشابكة "تشتمل على مئات البلايين من الكلمات من النصوص ويمكن استعمالها لجميع أنواع البحث اللغوي." (Adam Kigarriff and Greogroy Grefenstette (٢٠٠٣))

أما مصطلح لسانيات المدونات corpus linguistics فيطلق عادة على دراسة اللغة في ضوء النصوص اللغوية المدونة والمخزنة حاسوبيا. من ثم فهو ليس نظرية جديدة في اللسانيات بقدر ماهو منهج جديد أو مقاربة approach في البحث اللساني، كما سنبين أدناه.

١-٢- إنشاء المدونة اللغوية:

أما إذا نظرنا إلى المدونات اللغوية من المنظور اللساني الحديث، فنود الإشارة إلى أن من الأمور التي يجب أخذها بعين الاعتبار عند تصميم المدونة اللغوية وإنشائها ما يلي:

أ- تحديد الهدف من المدونة، مثلًا دراسة ظواهر لغوية معيّنة أو التعرُّف على الملامح المميّزة لنوع من النصوص، أولأسلوب كاتب معيَّن، أو التطور اللغوي وغير ذلك. (انظر: أنواع المدونات أعلاه.)

ب- تحديد النصوص المناسبة، من حيث النوع والزمان والمكان والمصدر. فإذا كانت الغاية إجراء دراسات على اللغة العربية المعاصرة في صورتها المكتوبة، فلا بد أن تشمل المدونة نصوصًا تمثل أنواع النصوص والأجناس الأدبية المختلفة من أدب (شعر، ونثر قصصي وغيرقصصي، مسرح، مقالات أدبية...) والتخصصات العلمية المختلفة وكتابات لمؤلفين مختلفين من سائر الدول العربية، وأعمال أنتجت في الفترة المحددة، وهكذا.

جـ- تحديد نسب كل نوع، وفق أسس علمية أو منهجية: ويخضع هذا لمعايير إحصائية (مثل نسبة الكتابات الصحفية والكتب العلمية المنتجة في فترة محددة إلى الأعمال الأخرى)، كما يستفاد في ذلك من آراء الخبراء والمتخصصين في مجالات المعرفة المخلفة.

د- البحث عن مصادر هذه النصوص (مكتوبة، مطبوعة، رقمية، صوتية أو سمعية-بصرية)

هـ- جمع النصوص من مصادرها المختلفة (انظر "وسائل جمع النصوص" أدناه.)

و- تحويلها إلى نصوص رقمية موحدة من حيث التشفير. (انظر "وسائل جمع النصوص" أدناه.)

ز- وسم tagging النصوص (العنونة الصرفية والنحوية – كما يسميها صوالحة –أي تحديد قسم الكلام لكل لفظة مثلا)، إذا أمكن. جدير بالذكر أن بعض الباحثين يستخدمون مصطلح التحشية annotation بدلا من الوسم. (انظر أدناه الوسم الصرفي

والوسم النحوي.) يشرح ليتش (Leech, ٢٠٠٥) التحشية بقوله: "تحشية النص هي إضافة معلومات تفسيرية للمدونة. مثلًا، من الممارسات الشائعة للتحشية إضافة الوسوم tags أو العناوين labels لبيان أقسام الكلام التي تنتمي إليها كلمات النص."

وللغوي الإنجليزي جوفري ليتش دراسة مستفيضة في هذا الموضوع، يناقشه من زوايا متعددة (انظر Leech, ٢٠٠٤).

في جميع الأحوال لابد من توافر حد أدنى للتعليقات أو التعليم markup على كل نص (المصدر، التاريخ، المؤلف مثلا).

وتقدم لنا لطيفة السليطي وإريك أتويل دراسة مفيدة في تصميم مدونة للغة العربية المعاصرة. ((Al-Sulaiti, L and Atwell, E) ٢٠٠٦).

١-٢-١. وسائل جمع نصوص المدونة وتخزينها:

من المعروف أن أية مدونة لغوية لا بد منها من جمع نصوصها من مصادر مختلفة وتخزينها وفق معايير محددة، تتفق مع الأهداف التي تنشأ المدونة من أجلها. ويختلف هذا الجمع تبعًا لنوع النص وطبيعته ومصدره. ونورد أدناه الأساليب المختلفة لهذا الجمع.

أ- النصوص الشفوية:

هناك عدة طرائق لتخزين النصوص الشفوية حاسوبيا: (١) تحويل النص إلى نص مكتوب باستخدام الإملاء التقليدي. ويمكننا استخدام برامج الإملاء الآلي لهذا الغرض، مع ضرورة المراجعة الدقيقة لها. (٢) كتابة النص كتابة صوتية، أي باستخدام رموز بديلة للألفباء الصوتية phonetic alphabet. وهذا مطلوب في تخزين النصوص العامية أو اللهجية وكذلك لغة الطفل. (٣) الكتابة الصوتية مع إضافة رموز للدلالة على الوقفة القصيرة والطويلة والتردد وماشابه ذلك من ظواهر يتسم بها الكلام المنطوق والحوارات. وفي جميع الأحوال لا بد لنا من تحديد المتحدث في الحوار.(لمزيد من التفاصيل بخصوص كتابة النصوص الشفوية، انظر Thomson, P., ٢٠٠٤).

ب- النصوص المكتوبة يدويا:

الوسيلة المعروفة هي تخزين النصوص يدويا عن طريق لوحة المفاتيح. ولكن يمكننا تخزين هذه النصوص نظريا على الأقل عن طريق الإملاء الآلي، بقراءة النصوص واستعمال نظام تحويل النص الشفوي إلى نص مكتوب (كلام إلى نص speech-to-text). وقد ظهرت في الآونة الأخيرة تطبيقات تحول الخط اليدوي إلى مادة مطبوعة رقمية.

جـ- النصوص المطبوعة:

هناك وسيلتان لتخزين هذه النصوص حاسوبيا: الأسلوب التقليدي هو استخدام لوحة الطباعة/ المفاتيح في الحاسوب. أما الأسلوب الآخر فهو باستخدام قارئة المحارف البصرية ((optical character reader (OCR)، حيث يقوم البرنامج بمسح النص المطبوع ثم تحويله إلى نص رقمي. غيرأن هناك مشكلات تواجه هذه الوسيلة: (١) محدودية برامج قراءة المحارف البصرية

من حيث تعاملها مع أبناط معينة (أي أشكال محدودة من الحروف)، (٢) ضرورة الوضوح التام لحروف النص المطبوع، (٣) وجود نسبة تقل أو تكثر من الخطأ في عملية قراءة المحارف، مما يتطلب تدقيقا يدويًّا.

(من أمثلة هذا النوع من البرمجيات "القارئ الآلي" من إنتاج شركة صخر المعروفة، والذي يتعامل مع الحروف العربية واللاتينية.)

د- النصوص الرقمية:

هناك عدة أنواع من النصوص المتاحة رقميا، لعل أهمها ما يلي:

١- النصوص الناتجة عن استخدام برمجيات معالجة النصوص/ الكلمات word processors.

٢- النصوص المتاحة على الشابكة (الإنترنت).

٣- النصوص المتاحة في صورة PDF.

٤- المواد المتاحة لدى شركات الطباعة والناشرين.

وتحتاج كل هذه النصوص إلى (١) تحويلها إلى نصوص خالية من المعلومات الخاصة بالشكل format، أي نصوص رقمية بسيطة plain text ،(٢) مع مراعاة توحيد الشفرة code المستخدمة، لكي تتعامل معها برامج تحليل النصوص، مثل المكشاف السياقي concordancer.)٣) كذلك يلاحظ ضرورة تحويل النص المتاح في صورة PDF إلى نص من الحروف قابل للتعديل عليه، من خلال مايسمى بمحولة البي دي إف PDF converter، ثم إخضاعه للخطوتين الأولى والثانية السابقتين.

٧- مصطلحات مهمة في مجال لسانيات المدونات:

نورد أدناه بعضا من أهم مصطلحات هذا العلم. ولمزيد من هذه المصطلحات، ينظر"مصطلحات لسانيات المدونات اللغوية" في: مدونة محمود إسماعيل صالح على الشابكة (الإنترنت). كما سيرد مسرد مصطلحات في نهاية هذا الكتاب.

أ- تعليم المدونة markup: إعطاء معلومات عن مصدر المدونة وتاريخ النص ومؤلف النص ونوع النص genre ويشمل الجنس الأدبي.ويكون ذلك خارج النص. لكن يلاحظ أن بعض الباحثين يستعملون المصطلح ليشمل جميع أنواع المعلومات التي تضاف إلى النص سواء داخله أو خارجه، مما يعنى اشتماله على التحشية والوسم اللذين سيرد التعريف بهما.

ب- التحشية annotation: إضافة معلومات في داخل النص، مثل االوسم tagging. وهناك أنواع مختلفة للوسم، منها:

(١) قسم الكلام (اسم، فعل، حرف). ويجدر ملاحظة أن التقسيم العربي التقليدي يحتاج إلى مراجعة وتفصيل أكثر عند وسم الكلمات، حتى يكون التحليل أكثر فائدة ودقة، بناء على أهداف المدونات التي تختلف عن أهداف واضعي النحو وتقسيماتها. (انظر: محمود إسماعيل الصيني (١٤١٢) "نحو معجم عربي للتطبيقات الحاسوبية"، وانظر تطبيقا عمليا لنوع جديد من التقسيم

في المعجم الذي أعده بكوولتر وباركنسون (٢٠١١) Buckwalter and Parkinson، وهما من رواد العمل في مجال مدونات اللغة العربية في الولايات المتحدة الأمريكية.)

يجدر بالذكر أن كتب النحو العربي تزخر بمصطلحات تتعلق بتصنيف الكلمات العربية يمكننا الإفادة منها، من مثل: ضمير، اسم موصول، اسم جامد، اسم مشتق(اسم فاعل، اسم مفعول، صيغة مبالغة، صفة مشبهة، اسم آلة، اسم مكان)، اسم جمع، اسم نوع، اسم علم، مصدر،ظرف (مكان، زمان)، فعل (لازم، متعد لمفعول، متعد لمفعولين، ناقص، ماض، مضارع، أمر...) ، حرف (جر، نصب، جزم، عطف، نفي...). ويجب وسم الكلمات بمثل هذه التصنيفات لإجراء كثير من الدراسات المعجمية والنحوية. (انظر مجدي صوالحة "العنونة الصرفية والنحوية".

. انظر أيضا (٢٠١٠) Kais Dukes and Nizar Habash، وكذلك مقال صوالحة وأتويل (٢٠١١) Sawalha and Atwell

(٢) الوسم النحوي أو الإعرابي (لبيان الوظيفة النحوية / الحالة الإعرابية):

مثالان للوسم الإعرابي:

الجملة الأولى: "الإيمان نور القلوب"

الوسم الإعرابي: (مب (الإيمان) خب (مض (نور) مض إل (القلوب)))

الشرح:

الإيمان = مبتدأ (مب)

نور القلوب = خبر (خب)

نور= مضاف (مض)

القلوب= مضاف إليه (مض إل)

المستوى الأول: الإيمان (مبتدأ) + نور القلوب (خبر)

المستوى الثاني: نور (مضاف) + القلوب (مضاف إليه)

(يلاحظ أن عدد الأقواس يشير إلى مستوى التحليل.)

الجملة الثانية: "الإيمان ينير القلوب"

الوسم الإعرابي: (مب (الإيمان) خب (فع (ينير) مف (القلوب)))

الإيمان = مبتدأ

ينير القلوب = خبر

ينير = فعل

القلوب = مفعول به

المستوى الأول: الإيمان (مبتدأ) + ينير القلوب (خبر)

المستوى الثاني: ينير (فعل) + القلوب (مفعول به)

وجدير بالذكر أن المصطلح بنك الشجرات treebank يستعمل للمدونة الموسومة إعرابيا. وقد يرجع سبب هذه التسمية إلى ما يسمى "التحليل الشجري" الذي يمثل التركيب بشجرة مقلوبة تتشعب فروعها ثنائيا غالبا (مثلا تبدأ بالجملة، ثم المسند والمسند إليه، ثم مكونات كل منهما، وهكذا).

ومن يرغب في الاطلاع على مزيد من المعلومات والأمثلة المتعلقة باللغة العربية، يمكنه الرجوع إلى الدراسة التي تمت في جامعة ليدز البريطانية (انظر: Kais Dukes, Eric Atwell and Abdul-Baqee Sharif (٢٠١٠,)، كذلك الدراسة التي يقوم بها فريق من الباحثين في جامعة كولومبيا الأمريكية Habash, Nizar, Reem Faraj and Ryan Roth).

(٣) الوسم الدلالي: المحددات الدلالية: يقصد بها تحديد المعنى المقصود من كلمة ما بصورة مختصرة، وهو أمر لابد منه خاصة للمشتركات اللفظية. فكلمة "عين" قد تأتي للدلالة على العين الباصرة (للإنسان عينان)، وعين الماء (ورد القوم العين ليستقوا)، وعين بمعنى الجاسوس (بث القائد العيون ليرصدوا تحركات العدو) وعين التوكيد في النحو (هذه القضية عينها).

(للمزيد من المعلومات،انظر: الملحق (ب) "معجم مصطلحات المدونات اللغوية".

٨- أساليب الوسم المختلفة:

أما طريقة الوسم فيستحسن اتباع مايعرف بلغة التعليم المعيارية العامة General Standard Markup Language (GSML). ولأمثلة تطبيقيةعلى اللغة العربية، انظر الدراسات المشار إليها أعلاه عن مشروعات جامعة ليدز البريطانية وجامعة كولومبيا الأمريكية ودرسات ديوكز وزملائه Dukes, Atwell and Sharif وحبش وزميليه Habash, Faraj and Roth.

ومن الأمثلة على أساليب الوسم المختلفة استخدام القوس المثلث <>قبل وبعد الجزء المطلوب وسمه. مثلا، لبيان أن النص من الشعر الأموي نكتب:< شعر أموي> أبيات من الشعر الأموي <شعر أموي/>، حيث تشير المعلومات بين القوسين المثلثين إلى أن النص من الشعر الأموي، كما تبين المعلومات بداية النص ونهايته. ومن الأمثلة الأخرى استخدام الأقواس المربعة، مثلا [S....S] للدلالة على بداية الجملة ونهايتها.

ولا بد من التأكيد على ضرورة بيان حدود العبارات التي تزيد عن كلمة واحدة للتعرف على بداية العبارة ونهايتها. وهذا ينطبق على التعبيرات الاصطلاحية والأمثال والأفعال المتعدية بحروف وماشابهها، إذا أردنا للحاسوب أن يعيننا في حصرها وإحصائها آليا.

ويمكننا وسم الكلمات بطريقة بسيطة، مثل: س-كتب، ف-كتب للتمييز بين الكلمة بوصفها اسما وبوصفها فعلا، وذلك في غياب برنامج للتحليل الصرفي.

ويذكر Bowker and Peason (٢٠٠٢) أن هناك نظاما يسمى معيار تشفير المدونة Corpus Encoding Standard (CES) اتفق عليه الباحثون في أوربا وأمريكا لتعليم النصوص وتحشيتها، تنقسم إلى ثلاث فئات هي:

(أ) التوثيق، ويقصد به الإشارة إلى معلومات ببليوغرافية عن النص ولغته، مثلا.

(ب) بيانات أساسية: وتشمل معلومات عن مكونات النص، مثل الفقرات والعناوين والهوامش...إلخ، إضافة إلى معلومات عن الرسومات graphics.

(جـ) تحشية لغوية: وتشمل المعلومات اللغوية داخل النص، مثل سمات الكلمات الخاصة بقسم الكلام part of speech والمعلومات الخاصة بالخطاب والوسم النحوي والوسم الدلالي...

جدير بالذكر أن مفهوم التعليم markup وكذلك نظام/ لغة ال SGML مطبقان بصورة معدلة على بعض النصوص المتوافرة على الشابكة internet، باسم HTML, XML. ويعني HTML (hypertext markup language) "لغة تعليم النص التشعبي" ويعني XML (extended markup language) لغة التعليم الموسعة". (انظر :Bowker and Peason (٢٠٠٢) ٨٠ -٨٩).

٩- أمثلة لمدونات للغة العربية

ذخيرة اللغة العربية:قبل أن نعطي وصفا لبعض المدونات، أود الإشارة إلى الدراسة القيمة لعبدالرحمن حاج صالح، أستاذ اللسانيات الجزائري الذي يعتبر بحق أول من دعى إلى إنشاء مدونة للغة العربية تسمى ذخيرة اللغة العربية منذ عدة عقود. وقد نشرت الدراسة في مجلة اللسان العربي التابع لمكتب تنسيق التعريب في العدد ٤٨ (١٩٩٩). ونورد أدناه أهداف المشروع ومواصفاته كما وردت في الدراسة المذكورة:

"أ- أهداف المشروع:

يرمي مشروع الذخيرة اللغوية العربية إلى إنجاز ما يلي:

١. بنك آلي للغة العربية المستعملة بالفعل (بنك نصوص)

٢. معجم آلي جامع للغة العربية مع المقابل الفرنسي والإنكليزي يستخرج من البنك الألي المذكور (معجم مفردات).

"ب – مواصفات المشروع:

سينجز البنك الآلي (أو الحاسوبي) للمعطيات النصية انطلاقا من الاستعمال الحقيقي للغة العربية ليضم:

١- المؤلفات ذات القيمة الكبيرة في الآداب والعلوم والتكنولوجيا وغيرها، القديمة منها والحديثة،

٢- المحاضرات الجامعية القيمة المنشورة،

٣- المقالات ذات القيمة المنشورة في المجلات الأدبية والعلمية والبحوث القيمة المعروضة في الندوات والمؤتمرات والموائد المستديرة وغيرها،

٤- جميع المعاجم العربية والمزدوجة اللغة القديمة والحديثة (مثل لسان العرب والمعجم الكبير الحديث وغيرهما). والغرض من بنك النصوص الآلي هو أن يكون قاعدة معطيات (أي بيانات) دائمة بحيث تقبل الزيادة والتصحيح على الدوام بسبب تطور المعلومات من خلال الاستعمال الحقيقي للغة العربية وبالتالي أن تصير المصدر الأساسي لإنجاز المعجم الجامع للغة العربية الذي سيحرره العلماء وخاصة أعضاء المجامع العربية وإنجاز العدد الكبير جدا من الدراسات والبحوث في اللغة العربية..." (عبدالرحمن الحاج صالح (١٩٩٩).

ولكن هذا المشروع الطموح لم يجد طريقه إلى التنفيذ حسب علمنا، لأسباب عديدة يتعلق بعضها بالمشروع وطبيعته وبعضها بعدم توافر الموارد البشرية والمادية اللازمة.

٨-١- مدونة معجم الطلاب:

وهي مدونة أعدها لجنة من اللغويين السعوديين تشمل محمود إسماعيل صالح، و علي الخبتي وعبد الرحمن الشمري و عبدالله الشلال. وكان الهدف منها تأليف معجم للتلميذ في المرحلتين المتوسطة والثانوية.

وقد كانت مكونات المدونة:

- المقررات الدراسية.
- القرآن الكريم
- الحديث النبوي الشريف (رياض الصالحين)
- نماذج أدبية عامة
- صحف ومجلات شباب
- برامج إذاعية ومتلفزة شبابية
- كتابات الطلاب. نصوص المدونة مصنفة (حسب النوع والمصدر والمستوى)

ومجموع كلمات المدونة ثلاثة ملايين كلمة.

٨-٢- المدونة اللغوية العربية (المدونة العربية):

لاشك أن هذه المدونة تعتبر أكبر مدونة مفتوحة للغة العربية وأكثرها تنوعا من حيث النصوص والمصادر.

نورد أدناه مقتطفات من النبذة الرسمية عن المدونة من موقعها:

- تصميم المدونة العربية:

إن مدونة مدينة الملك عبدالعزيز للعلوم والتقنية، ستكون أكبر وأضخم مدونة لغوية للعربية. وهي في مرحلتها الاولى تسعى لجمع سبعمائة مليون كلمه وسوف يزداد حجمها إلى أن تصل الى بليون كلمة في مرحلة لاحقة ان شاء الله. راعى تصميم المدونة اللغوية العربية لمدينة الملك عبدالعزيز للعلوم والتقنية عدة معايير خارجية لاختيار نصوص المدونة تعتمد على خمس ركائز أساسية هي: البعد الزمني، والبعد الجغرافي، والوعاء المعلوماتي، و المجال المعرفي والتصنيف الموضوعي. إضافة إلى هذا فإن المدونة في مرحلتها الحالية هي للنصوص المكتوبة والكاملة فقط ولا تحوي أي نصوص منطوقه مثل الحوارات التلفزيونية أو الخطابات السياسية أو أي نصوص غير مكتملة مثل فصل من كتاب أو جزء من مقال.

- البعد الزمني

أول المعايير التي تم مراعاتها هو عامل الزمن أو البعد التاريخي والذي يمتد من عصر ما قبل الإسلام وحتى عصرنا هذا والذي أثر على الوعاء أو الصورة التي ظهر فيها النص وكذلك على حجم النصوص المطلوب جمعها لكل فترة زمنية والذي كان على هيئة دالة أسية تتناسب مع تطور المعارف والعلوم والتدوين الخاص باللغة العربية بحيث يزداد عدد كلمات المدونة كلما قربنا من العصر الحديث ويزداد تنوع صورها معه كذلك.

- البعد الجغرافي

ويقصد به المكان الذي صدر منه النص. ولأن المدونة تعنى باللغة العربية بمجموعها وتحاول أن تكون ممثلة للغة ومتغيراتها فإنه لم يتم تحديد بلد عربي بعينه لجمع النصوص بل إن المطلوب هو تنوع البلدان والكّتاب أو المؤلفين. ويسعى تصميم المدونة لجمع النصوص من جميع البلدان والمناطق العربية قديما وحديثا، كما أن التصميم لا يمنع ضم أي نص مكتوب باللغة العربية من أي بلد كان بحيث لا يطغى بلد او منطقة على بقية البلدان او المناطق قدر الإمكان.

- أوعية النشر:

هذا هو البعد الثالث لتصميم المدونة. حيث تم اختيار ما يناسب كل فترة زمنية وما كان سائدا فيها من علوم ومعارف وما كان أكثر انتشارا وتداولا بين الناس من أوعية للنشر وما تكون لغته مناسبة ورصينة. فمثلا تم استبعاد المنتديات الحوارية و صفحات الانترنت الخاصة والتي يغلب عليها هذا الوقت اللهجات الدارجة ولا تتقيد باللغة العربية الفصحى. وتم اختيار عشرة أوعية للنشر وهي المخطوطات المحققة، الصحف، المجلات، الكتب، الرسائل الجامعية، الدوريات المحكمة، الإصدارات الرسمية،

وكالات الأنباء، الإنترنت والمناهج الدراسية. وتم اختيار هذه الأوعية بناء على انتشارها وتأثيرها و رصانة لغتها. وكل فترة زمنية من فترات المدونة لها ما يناسبها من هذه الأوعية.

- المجال المعرفي:

يندرج تحت كل وعاء من الأوعية المختارة مجالات مناسبة له، تحدد مجال النص وسمته العامة. ففي الصحف على سبيل المثال هناك مجالان رئيسان هما الأخبار والمقالات. وفي المخطوطات المحققة وفي الفترة التي كتبت فيها هذه المخطوطات كان هناك مجالات عامة مثل العقائد والفقه وأصوله وعلوم اللغة وغيرها بما يناسب كل فترة. وينطبق هذا على كل وعاء من الأوعية. وهذا البعد يعطي المدونة فرصة أكبر لإيضاح الاختلافات بين كل مجال وآخر وفترة وأخرى كما يوضح أيضا تنوعها وتمثيلها للغة بشكل أكبر.

- التصنيف الموضوعي

يندرج تحت كل مجال من المجالات المخصصة للأوعية عدة مواضيع تفصل المجال وتوضح تنوعاته الأدق وتظهر الثراء والتنوع في كل مجال ووعاء. ففي وعاء الصحف وتحت مجال الأخبار هناك عدة مواضيع مثل الأخبار الاجتماعية، الاخبار السياسية، الأخبار الرياضية، الأخبار الاقتصادية،.....الخ. وفي وعاء المخطوطات المحققة وتحت مجال أصول الفقه هناك عدة مواضيع مثل أصول الفقه الشافعي، أصول الفقه الحنبلي، أصول الفقه المالكي، أصول الفقه الحنفي وأصول الفقه الإثنا عشري. يعطي هذا البعد تنوعا وثراء وتخصصا أكثر لنصوص المدونة مما يجعل الفائدة أكبر منها بحيث يجد الدارس ما يناسبه لاختياره ودراسته ومقارنته.

معلومات عامة عن المدونة:

عدد الكلمات الكلي = ٧٣٩,١١٩,٠١١

عدد الكلمات بدون تكرار = ٧,٤٦٤,٣٩٦

العدد الكلي للنصوص = ٩٥٠,٤٧٨نصّا

العدد الكلي للمؤلفين = ١,٩٠٠ مؤلفاً

(المصدر: الموقع الرسمي للمدونة اللغوية العربية لمدينة الملك عبد العزيز للعلوم والتقنية www.kacst.org.sa).

يجدر بنا الإشارة إلى أن هذه المعلومات كانت منشورة قبل العام ٢٠١٥م. فالحجم الحالي للمدونة يتجاوز المليار كلمة (عبدالمحسن الثبيتي – تواصل شفوي).

يجدر بالذكر أن المدينة طورت مؤخرا برنامجا لتحليل المدونات اللغوية باسم "الغواص"، وهو متاح على الشابكة (الإنترنت): http//sourceforge.net/projects/kacst-acptool

٧-٣- المدونة اللغوية العربية الدولية:

من مشروعات المدونات العربية الطموحة المدونة العربية الدولية International Corpus of Arabic التابع لمكتبة الإسكندرية، وهي مدونة للعربية المعاصرة تطمح إلى بناء مدونة قوامها مائة مليون كلمة. (انظر Al-Ansary, Sameh, Nagi, Magdy and Adly, Nohaفي قائمة المراجع.)

ونورد أدناه مقتبسات من التعريف الرسمي للمدونة المذكورة:

- **هدف المدونة اللغوية العربية العالمية**

بناء مدونة لغوية للعربية المعاصرة تحوي ١٠٠ مليون كلمة محللة صرفيا ونحويا ودلاليا، وقد روعي فيها أن تكون ممثلة لقطاع إقليمي كبير من الدول الناطقة باللغة العربية المعاصرة وعاكسة بشكل حقيقي وواقعي لأنماط استخدام اللغة العربية المعاصرة في أنحاء العالم العربي. بمجرد الانتهاء من بناء المدونة ستكون أول مدونة محللة ومتاحة كمورد لغوي للباحثين بصفة عامة والباحثين اللغويين بصفة خاصة لتفيد في وصف نظريات اللغة من خلال الاستخدام الواقعي للكلمات.

- **تخطيط المدونة اللغوية العربية العالمية**

لقد روعيت العديد من الأمور المرتبطة ببناء المدونة مثل التمثيل الجيد للنصوص في العربية المعاصرة والتنوع في فئات النصوص ومحتواها والتوازن بين كل فئة من النصوص وحجم الكلمات المجمعة في كل فئة من فئات التجميع. عند النظر إلى تمثيل العربية المعاصرة داخل المدونة نجد أن الاهتمام الأساسي هو التغطية والتمثيل الواقعي لمختلف المصادر من كل المجتمعات العربية. فشملت المدونة عددا من المصادر والفئات المختلفة للنصوص وذلك بهدف تحقيق شروط التمثيل الجيد ومدى انتشار المصدر أو الفئة، والتوازن بين كل مصدر وكل فئة، وحجم الكلمات في كل مصدر وفئة.

ونجد أن تصميم المدونة اعتمد بالأساس على البدء بحصر المصادر المختلفة، وداخل كل مصدر تم إدراج الفئات المميزة له. وقد تم حفظ النصوص داخل المدونة بطريقة هرمية من خلال تسمية النصوص بطريقة توضح العديد من المعلومات مثل المصدر والفئة وتاريخ النشر.

لقد تم الأخذ في الاعتبار العديد من الأمور عند تجميع المدونة مثل عدد الفئات المتضمنة داخل المدونة، وعدد النصوص داخل كل فئة من هذه الفئات، بالإضافة إلى متوسط عدد الكلمات داخل كل نص. تبعا لطبيعة كل مصدر من مصادر التجميع.

- **تصميم المدونة اللغوية العربية العالمية**

- يوجد أربعة مصادر أساسية: الصحافة والمقالات الإلكترونية والكتب والدراسات الأكاديمية.

المصدر الخاص بالصحافة منقسم إلى ثلاثة مصادر فرعية: الجرائد والمجلات والصحافة الإلكترونية.

- يوجد إحدى عشرة فئة على مستوي المدونة: العلوم الإستراتيجية والعلوم الاجتماعية والرياضة والدين والأدب والعلوم الإنسانية والعلوم الطبيعية والعلوم التطبيقية والفنون والثقافة والسير الذاتية والنصوص المتنوعة.
- يوجد أربع وعشرون فئة فرعية: سياسة وقانون واقتصاد واجتماع ودين إسلامي ودين مسيحي وأديان أخرى ودين مقارن وقصص وشعر ونثر ودراسات لغوية وأدبية وطب وهندسة وزراعة وتكنولوجيا وعلم الأحياء وعلم الفيزياء وعلم الفضاء وعلم الجيولوجيا والبيئة وعلم الكيمياء وعلم النفس وعلم الفلسفة وتاريخ.
- يوجد أربع فئات فرعية من فئة القصص الفرعية: روايات وقصص قصيرة وقصص أطفال ومسرحيات.
- تغطي المدونة جميع المنشورات داخل الوطن العربي وكذلك بعض المنشورات العربية المنشورة خارج الوطن العربي.

- **تحليل المدونة اللغوية العربية العالمية**

تشمل هذه المرحلة حاليا التحليل الصرفي لكل كلمة موجودة داخل المدونة، وقد تم في هذه المرحلة تحليل المدونة بطريقة آلية مبنية على بعض الطرق الإحصائية وبعض القواعد اللغوية بالاعتماد على أحد المحللات الصرفية الشهيرة – تيم باك والتر (Tim Buckwalter) – حيث يوضح التحليل الصرفي عدد من المعلومات كالسوابق واللواحق قسم الكلمة وساقها وجذعها وجذرها ووزنها الصرفي بالإضافة إلى نوع الكلمة من حيث الجنس والعدد والتعريف تبعا للسياقات المختلفة للكلمات داخل كل نص."

(المصدر: المدونة العربية الدولية International Corpus of Arabic التابع لمكتبة الاسكندرية).

٧-٤- بعض مدونات اللغة العربية الأخرى:

إن أفضل حصر للمدونات العربية إلى عهد قريب هو ما قامت به الباحثة لطيفة السليطي من جامعة ليدز Al-Sulaitie, Latifa(٢٠١٠). Arabic Corpora. غير أن القائمة مكتوبة باللغة الإنجليزية.

ونورد في الملحق (أ) أسماء المدونات اللغوية العربية المشهورة، وتشمل ترجمة لقائمة لطيفة السليطي، متبوعة بقائمتين إضافيتين نلحقهما بالقائمة المذكورة، وتشمل كذلك عددا من المدونات الموسومة.

١٠- متطلبات التعامل مع المدونات:

تتطلب الإفادة المثلى من المدونات مايلي:

- محرك بحث search engine: وهذا أبسط برنامج يفيد الباحث في العثور على الكلمات في سياقات مختلفة. وينبغي التنبيه إلى أن البرنامج قد يكون بسيطا بحيث ينظرإلى الكلمة بمعناها الحاسوبي (حسب الشكل فقط)، وهو البرنامج المتوافر حاليا مع "مدونة اللغة العربية" وبرنامج "غواص" التابعين لمدينة الملك عبدالعزيز للعلوم والتقنية بالرياض، أو يكون متطورا بحيث يشمل تحليلا وتركيبا صرفيا، حيث يمكن البحث بالجذر مثلا أو بالجذع، ويورد البرنامج الكلمة في صورها المختلفة. وهذا النوع الأخير متوافر مع مدونة اللغة العربية التابعة لجامعة بريغهام يونغ Brigham Young University (BYU) في ولاية يوتا الأمريكية. (انظر:ArabiCorpus: arabic corpus search tool).

- برنامج المكشاف السياقي concordancer ، لإعداد الكشاف السياقي – أي قائمة بألفاظ النص/ المدونة في سياقاتها (مايسمى بالإنجليزية Key Word in Context (KWIC)، بترتيبات مختلفة:

١- الأصل: الكلمة المفتاحية key word مسبوقة ومتبوعة بعدد من الكلمات، حسب ورودها في المدونة. ولكن يمكننا أن نطلب من البرنامج أن:

٢- يجعل ترتيب أسطر الكشاف بناء على الكلمةالمفتاحية وما يسبقها مباشرة <...قرأت **كتابا**....،...عندي **كتاب**...>. ويسمى الترتيب الموجه يمينا right sorted

٣- أو بناء على الكلمةالمفتاحية ثم ما يتبعها مباشرة <... **كتاب** جديد...، **كتاب** يتحدث عن...> ويسمى الترتيب الموجه يسارا left sorted. ويجدر بالذكرأن ربط الكلمتين "يمين" و"يسار" بالسابق واللاحق يختلف في العربية عن المفهوم في اللغات الأوربية التي تتجه الكتابة فيها من اليسار إلى اليمين.

هذا ويمتاز المكشاف السياقي بإمكانات مهمة كثيرة، يجعله يختلف عن محرك البحث التقليدي، مثل إحصاء التكرار وترتيب كلمات المدونة وفق شيوعها (تنازليا أو تصاعديا، أي بدءا بالأشيع أو بالأقل شيوعا)، إضافة إلى العمل الأساسي لها، وهو إيراد الكلمات في سياقاتها.

ومن أمثلة المكشاف السياقي برنامج aconcorde الذي أعده مجموعة من الباحثين في جامعة ليدز البريطانية، ويعمل مع النصوص العربية والإنجليزية وبرنامج "الغواص Ghawwas اللذان صمما للباحث العربي، وبرنامجا Monoconc, Sketch Engine اللذان يعملان بسلاسة نسبية مع النصوص العربية، غير أن تعليماتهما باللغة الإنجليزية . وبرنامج Word Smith، وهو برنامج متطور من إعداد مطبعة جامعة أكسفورد Oxford University Press، ولكنه لا يعمل بسلاسة مع النصوص العربية.

- برنامج محلل صرفي morphological analyzer (للتعرف على الصورة الأساسية للكلمة – الجذع- وكذلك الجذر والوزن)، وكذلك للفصل بين السوابق واللواحق المتصلة بالكلمات، مثل بعض حروف الجر والعطف المتصلة بالكلمة، وغير ذلك. مثلا، كلمة "فهم" قد تكون ف+هُم (الضمير)، أو ف+هَمَّ (الفعل)، أوالفعل فَهِمَ، أو الاسم فَهمٌ. و الفعل "يكتبون" تعاد إلى جذعها: كَتَبَ، إو إلى جذرها: ك ت ب والوزن فَعَلَ. هناك عدد من هذه البرامج يرد بعضها تحت مسمى "المجذّع stemmer" الذي يقوم بتجريد الكلمة العربية من للواصق (مثل حروف الجر والضمائر المتصلة، وإعادة الأفعال إلى صورها الأبسط لها – فعل ماض فاعله مفرد مذكر).

- برنامج لتحديد قسم الكلام part of speech الذي تنتمي إليه الكلمات، ما يسمى بالإنجليزية POS tagger. في غياب ذلك يحتاج الباحث إلى عمل ذلك يدويا، إذا كانت هناك حاجة لذلك. ومثال ذلك كلمة "كاتب" قد تكون اسم فاعل أو فعلا ماضيا. وكلمة "عين" قد تكون اسمًا مفردًا، أو جمعًا (كما في الحور العين)، أوفعلًا (عَيَّنَ فلانًا...) فلا بد من وسم كل كلمة لبيان قسم الكلام الذي تنتمي إليه. وهو أمر مطلوب في كثير من البحوث اللغوية. وهناك حاليًا عدد من المدونات العربية الموسومة نوردها في قائمة المدونات الموسومة.

- برنامج تشكيل آلي (يعتمد على الإعراب parsing اي تحديد وظيفة الكلمة النحوية). ويسمى برنامج الإعراب parser (أي المُعرِب).

جدير بالذكر أن للباحثة لطيفة السليطي دراسة مسحية ممتازة لأنواع البرمجيات المتاحة للتعامل مع اللغة العربية على مستوى المكشاف والتحليل الصرفي والتشكيل الآلي.(انظرwww.leeds.ac.uk/latifa/survey.htm)

ومن الأعمال الجديرة بالاطلاع والإفادة الدراسة التي نشرها مروان البواب (٢٠١٢) بعنوان "محركات البحث في النصوص العربية "، ففيه مناقشة مستفيضة للجوانب المهمة في محركات البحث في النصوص العربية ومشكلاتها.

هناك أيضا دراسة لصوالحة وأتويل يجريان فيها مقارنة بين بعض برامج التحليل الصرفي للغة العربية. (انظر Sawalha, M and Atwell, E ((٢٠٠٨).

- الوسم اليدوي manual tagging لإعطاء معلومات مختلفة عن الكلمات والتراكيب والجمل التي يحتاج إليها الباحث (مثل حدود التعبيرة الاصطلاحية أوالعبارات المسكوكة (كما يسميها البعض)، أي وضع علامة تدل على بداية العبارة وأخرى للدلالة على نهاية العبارة. ويلاحظ أن معظم المعلومات التدولية والخطاب تحتاج إلى وسم يدوي، فمعظم هذه المعلومات لايمكن القيام به آليا لعدم وجود برمجيات يمكنها القيام بذلك دون تدخل بشري.

- تطوير برنامج للوسم الآلي:

هناك برنامج مفتوح المصدر open source يسمى بوابة gate لمساعدة الباحث اللغوي في إعداد برمجيات للوسم الآلي. (الموقع: www.gate.ac.uk). يصفه الموقع بأنه:

"مصدر مفتوح لإيجاد حل دورة حياة كاملة لمعالجة النصوص a full-lifecycle open source solution for text processing".

١١- مجالات الإفادة من المدونات اللغوية:

١١-١- بعض أنواع التحليل في لسانيات المدونات:

(أ) إيجاد الترابط بين جوانب لغوية ولغوية أخرى:

١- معجمي-معجمي: أي كلمة مع مصاحبات لفظية، مثل باقة ورد، وقطيع ماشية، مواء الهر.

٢- معجمي-تركيبي: أي كلمة تتطلب تركيبًا ما...، مثلا: الفعل الذي ينصب مفعولين أصلهما مبتدأ وخبر أو ليس أصلهما مبتدأ وخبر، والفعل الذي يتبعه مصدر مؤول أو أن+فعل...

٣- تركيبي-تركيبي: أي تركيب يتطلب تركيبا معينا، مثل الخبر شبه الجملة الذي يتطلب وقوعه مقدما على المبتدأ، كما في: "في الدار عائلة كبيرة").

(ب) إيجاد ترابط لغوي-غير لغوي: أي العلاقة بين ألفاظ وتراكيب وأساليب لغوية ترتبط بمؤلف، أومنطقة جغرافية أو بلغة ولهجة ما، او ظواهر لغوية ترتبط بجنس أدبي (شعر أونثر مثلا)، أوظواهر لغوية (كالجمل الاسمية أوالبناء للمجهول) وارتباطها أو كثرة ورودها في تخصص معين (الطب أو الهندسة أو النقد)...

من أمثلة الترابط اللغوي-غير اللغوي ارتباط بعض الألفاظ بمنطقة جغرافية ما (كلمة "شغل" في المغرب العربي، مقابل "عمل" في المشرق العربي)، وعبارات ورّاق وسوق الورّاقين وارتباطها بلغة التراث العربي.

(جـ) دراسة أنواع النصوص المختلفة: أي دراسة النصوص الأدبية والعلمية والصحفية والقانونية مثلا وخصائص كل منها.

ويمكننا أن نصنف أسلوب التحليل إلى نوعين:

(١) تحليل كمي quantitative (إحصائي): مثلا، بالنسبة لكلمة ما، عدد مرات ورودها في النص، ما المصاحبات اللفظية لها وما شيوع كل منها.

(٢) تحليل كيفي/ نوعي qualitative: ويعني ذلك محاولة تفسير الظاهرة اللغوية، مثلا "لماذا تشيع كلمة معينة أو تركيب ما في نوع من النصوص؟"، مثل شيوع الضمائر في اللغة المحكية مقارنة باللغة المكتوبة أو شيوع المبني للمجهول في النصوص العلمية والتقنية (الإنجليزية)، مقارنة بالنصوص الأدبية مثلا. (انظر Biber et al., ١٩٩٨).

٩-٢- أهم مجالات الإفادة من لسانيات المدونات:

لا شك أن المدونات اللغوية تعتبر مصدرًا غنيًّا للبحث اللغوي بشتى أنواعه، بل مصدرًا أساسا له في كثيرمن الأحيان. وكما أشرنا في مكان سابق من هذاالفصل كان للمدونات المكتوبة والمحفوظة (الشفوية) دورها في استنباط قواعد كثير من اللغات وأنظمتها الصرفية والصوتية وأساليبها البلاغية ومعاجمها أو استقرائها، كما فعل علماء العربية واللسانيون الغربيون الذين درسوا اللغات الأفريقية والأسترالية ولغات الهنود الحمر. ومن خلال المدونات اللغوية (النصوص) توصل اللسانيون إلى معرفة العلاقات بين اللغات، ومن ثم تصنيفها إلى أسر وفروع، كما حدث في علم اللغة التاريخي.

ولعل من أوائل هذا النوع من الدراسات في العصر الحديث الأعمال المعجمية الإحصائية التي استفاد منها صناع المعاجم بأنواعها المختلفة ومعدو مواد تعليم اللغات، خاصة لغير الناطقين بها. ونورد أدناه نماذج للدراسات في مجالات مختلفة اعتمدت أو تعتمد على المدونات اللغوية.

٩-٢-١- الدراسات المعجمية وصناعة المعاجم **lexicography**:

كانت الدراسات المعجمية وصناعتها من أوائل التطبيقات العلمية والعملية للمدونات اللغوية (اليدوية في حينها)، حتى من قبل التدوين الإلكتروني أو الرقمي، كما هو واضح من أعمال مايكل وست (١٩٥٣) Michael West وثورندايك ولورج Thonrndike and Lorge,١٩٧٢.

ونجد شرحا جيدا لدور المدونات المحوسبة في كتابات سنكلير Sinclair الباحث الرئيس في مشروع مدونة COBUILD الذي تم بالتعاون بين جامعة بيرمنغهام Birmingham وناشر المعاجم كولينز Collins. (انظرمثلا: (١٩٩١) .Sinclair, J).

ولاشك أن هذا من أشيع نماذج الإفادة من المدونات اللغوية المطبوعة والمحوسبة. فقد كانت جميع قوائم شيوع الألفاظ في اللغات المختلفة مبنية على مدونات مختلفة. من هذه القوائم الرائدة في الإنلجليزية قوائم مايكل وست Michael West وثورندايك ولورج Thorndike and Lorge. وفي العربية قوائم بريل ولانداو وفاخر عاقل وداود عبده (انظر: داود عبده، ١٣٩٩هـ). فقد قام هؤلاء بالإحصاء اليدوي للكلمات في نصوص مطبوعة للتعرف على الكلمات الشائعة فيها. كذلك قام بعضهم بأبعد من ذلك، حيث قام مايكل وست بإحصاء تكرار المعاني المختلفة للألفاظ semantic count، كما هو واضح من عنوان كتابه (انظر Michael West, ١٩٥٣).

٩-٢-١-١- تحديد مفهوم الكلمة والوحدة المعجمية والمصطلحات ذات العلاقة:

من القضايا المهمة التي يجب التنبه لها عند الحديث عن الدراسات المعجمية التفريق بين المصطلحات ذات العلاقة بالكلمات وتعريفاتها المختلفة، ومنها ما يلي:

١- الكلمة الحاسوبية: مجموعة من الحروف (أوحتى رمز واحد) مسبوقة بفراغ ومتبوعة بفراغ، مثلا: الأرقام والفاصلة أو علامة الاستفهام إن كانت مسبوقة ومتبوعة بفراغات، والكلمات "ويكتبون"، "فكتبناهم" على الرغم من اشتمالها على عدة وحدات صرفية.

٢- الكلمة الفعلية / النصية token، وهي الكلمة كما ترد في النص بصور مختلفة (مثلا الفعل بتصريفاته المختلفة – كتب، كتبنا، كتبا، نكتب، يكتبون...أو الاسم في صيغة المفرد والمثني والجمع – تلميذ، تلميذان، تلاميذ...). ويستعمل المصطلح أحيانًا ليعني الكلمة الحاسوبية إن كانت لغوية (انظرأعلاه).

٣- الوحدة المعجمية lexeme: قد تكون كلمة واحدة أو عبارة تتكون من أكثر من كلمة، كما في التعبيرات الاصطلاحية (مثلا: ضرب أخماسا في أسداس، بمعنى احتار).

٤- االجذع stem (lemma) وهي الكلمة المجردة من الزوائد الصرفية كعلامات الجمع أو التثنية أو الدلالة على الزمن في الأفعال. وهو للاسم الكلمة في صيغة المفرد (وجمع التكسير، أحيانا)، مثل كاتب، مكتوب وللفعل صيغته في الماضي للمفرد المذكر، مثل:، كتب، قال. وهو ما نجده في مداخل المعاجم المرتبة حسب الكلمات (أو النطق).

٥- الكلمة النوعية type، وهي الجذع المشارإليه أعلاه.

٦- الجذر root (في العربية، مثل: ك ت ب، ق و ل) والمكون الأساس للكلمة في اللغات الأوربية (من أمثلة ذلك الجذور اليونانية واللاتينية، مثل bio, geo, logy)، إضافة الي الصورة الأبسط للكلمة.

٧- تجريد الكلمة من الزوائد (تجذيع lemmatization): ويعني ذلك في العربية تحويل الكلمة إلى صورتها الأساس (الجذع) مع حذف الزوائد السابقة، مثل حرف الجر "بـ" والعطف "فـ" واللاحقة مثل الضمائر المتصلة "ـه" و"ـها". فالكلمة الحاسوبية "فهم" تجرد إلى "فهم" الاسم والفعل، وكذلك إلى "فـ" +"هم". والكلمة الحاسوبية "بالقلم" نحذف منها الباء لتصبح القلم".

٨- قائمة الكلمات المستثناه stop list:

عند طلب الكشاف السياقي أو قوائم الشيوع في المدونة، هناك كلمات لانحتاج أحيانا إلى معرفة شيوعها (وهي ماتسمى بالكلمات الوظيفية، مثل الضمائر وحروف الجر أو حروف العطف)، لكثرتها ولأننا نعرف سلفا أنها أكثر الكلمات شيوعا في اللغة. في هذه الحالة يمكننا أن نعطي للبرنامج الإحصائي قائمة بهذه الألفاظ ليتجاهلها الحاسوب. وهذه هي التي تسمى قائمة الكلمات المستثناه stop list.

٩- شيوع الألفاظ word frequency:

للتعرف على مدى شيوع كلمة معينة أو تكرارها نقوم بعملية إحصاء الكلمات word count، مع ملاحظة التعريفات المختلفة لمفهوم الكلمة المذكورة أعلاه. والأصل في ذلك النظر إلى تكرار الكلمة النوعية أو الجذع بغض النظرعن صور ورودها فعليا (الكلمة الفعلية). ويجب التنبيه إلى عدم الخلط بين أنواع الكلمة بمفاهيمها المختلفة المذكورة بعاليه، وكذلك بين الكلمة والجذر.

فالجذر "غ ف ر" تشتق منه جذوع أو كلمات مختلفة تختلف من حيث شيوعها ومعانيها (غفر، استغفر، غافر، غفّار، غفور، مغفرة).

١٠- شيوع المعاني semantic count:

يشير هذا المصطلح إلى إحصاء المعاني المختلفة للكلمة متعددة المعاني polysemous، وينطبق ذلك خاصة على المشتركات اللفظية homonyms، مثل "عين" الباصرة ومنبع الماء وأداة التوكيد، فلا شك أن لكل من هذه المعاني تكرارها الخاص بها. ونجد خيرمثال لذلك في كتاب مايكل وست الريادي الذي أشرنا إليه أعلاه (Michael West, ١٩٥٣).

٩-٢-١-١- مجالات البحث التي تتعلق بالجانب المعجمي للغة:

من أهم هذه المجالات مايلي:

١- ما مدى شيوع كلمة ما، بصورة عامة أو في نصوص مختلفة النوع مثلا؟

٢- ما نسبة شيوع كلمة ما في نصين مختلفين؟

عند إجراء هذا النوع من المقارنة علينا أن نأخذ بعين الاعتبار حجم كل نص. لذلك يفرق الخبراء بين الإحصاء أو العد العام / الخام raw، أي ذكر التكرار في كل نص بغض النظر عن عدد كلماته. فالمفروض أن نلجأ إلى حساب التكرار النسبي normed count، أي نذكر تكرار الكلمة في كل ألف كلمة مثلا من النصين أ و ب، كأن نقول إن كلمة "مسألة" ترد ٥٠ مرة في كل ألف كلمة من النص أ وترد ٥٧ مرة في كل ألف كلمة من النص ب.

٣- ما مدى شيوع معاني كلمة ما (الإحصاء الدلالي semantic count) في نص ما أو في نصوص مختلفة ؟ نجد مثلا أنه في الاستعمال العام، ترد كلمة "عين" بمعنى البصر أكثر منها بمعنى عين الماء. كذلك نجد أن كلمة "باب" بمعنى المدخل قد يكون أكثر شيوعا منها بمعنى فصل من كتاب مثلا.

٤- هل للكلمة ارتباط بكلمات معينة (المصاحبات اللفظية – سابقة أو لاحقة)؟، مثلا: "جماعة من أهل العلم" و"فريق من اللاعبين" و"قطيع من الماشية".

٥- هل للكلمة ارتباط بلهجة أو نوع من الاستعمال اللغوي genre أو غيرذلك من العوامل غير اللغوية؟ (مثلا: كلمة تكوين بمعنى التدريب والشغل بمعنى العمل في المغرب العربي).

٦- ما مدى شيوع المصاحبات اللفظية المعينة، بصورة عامة أو في لهجات أو أنواع من الاستعمال اللغوي؟

٧- ما هي مجالات استعمال المترادفات (وسيم، جميل أو كبير، ضخم) وتوزيعاتها، حسب نوع النص مثلا ؟

٨- الاختلاف بين استعمالات الكلمة في حالتي الإفراد والجمع (مثلا كلمة "حبل" و"حبال" في القرآن الكريم، حيث ترد الأولى غالبًا بالمعنى المجازي بينما ترد الثانية بالمعني الحسي للحبل).

٩- التعرف على التعبيرات الاصطلاحية (المسكوكات) ومدى شيوعها عامة أو في أنواع مختلفة من النصوص. ويشمل ذلك الأفعال المتعدية بحروف. (انظرمثلا دراسة وفاء كامل (٢٠٠٧) ومحمد الحناش (١٩٩٦).

١٠-دراسة الجذوروالحروف العربية من حيث تكرارها وشيوعها. ويعتبر علي حلمي موسى في دراستيه (بالتعاون مع إبراهيم أنيس وعبدالصبور شاهين) والخاصتين بمعجمي **لسان العرب** و**الصحاح** بحق رائدا في هذا النوع من الدراسات (انظر قائمة المراجع). ومن الأعمال الرائدة في العالم العربي أيضا (المعجم الحاسوبي: إحصاء الأفعال العربية في المعجم العربي) من إعداد محمد المراياتي ويحي ميرعلم ومحمد حسن طيان. ولكن يلاحظ أن هذه الأعمال عاملت بعض المعاجم العربية بوصفها مدونات في إنجاز الدراسات المشار إليها.

(انظر الدراسة القيمة للدكتور عبدالرحمان حاج صالح، رائد مشروع الذخيرة العربية، في قائمة المراجع، حيث يذكر عددا من فوائد المدونة اللغوية الإضافية. ومن الكتب التي عالجت بإسهاب كثيرا من القضايا المعجمية كتاب سنكلير الذي كان من المشرفين على مشروع COBUILD ومحررًا لمعجم BBC للغة الإنجليزية: *Corpus, Concordance and Collocation*.)

٩-٢-١-١- صناعة المعاجم:

يلاحظ هنا أن إعداد المعاجم (أحادية اللغة وثنائيتها) يستخدم المدونات لأغراض مختلفة، مثل اختيار ألفاظ المعجم (المداخل) وكذلك حصرمعاني الألفاظ (بناء على الكشاف السياقي والنصوص الأصلية- الشواهد) وفي التمثيل لاستعمالات المداخل المختلفة (من واقع المدونة). من المشروعات الرائدة معجم *Collins COBUILD English Language Dictionary*. ومن المشروعات الجديدة معجم كامبريدج للمحتوى الأكاديمي *Cambridge Academic Content Dictionary* الذي صدر في عام ٢٠٠٩ والذي اعتمد مصادر رسمية وشبه رسمية للمعاييروالاختبارات المقننة وماشابه ذلك في اختيار مداخله.

كذلك يستخدم الحاسوب أحيانا في التحكم في لغة التعريف، كما فعلت لونغمان في معجمها *Longman Dictionary of Contemporary English*، حيث حددت هذه اللغة بالألفي كلمة الأكثر شيوعا في اللغة الإنجليزية.

نجد الشيء نفسه في معجم كامبريدج للمحتوى الأكاديمي المذكور آنفا والموجه للطلاب كما يبدو، حيث يذكر محرره أن "تعريفات المعجم كتبت باستخدام مفردات كامبريدج للتعريفات Cambridge Defining Vocabulary، وهي قائمة من حوالي ٢٥٠٠ كلمة الشائعة والتي يعرفها الطلاب. وقد تم تطوير هذه القائمة بالإفادة من الأجزاءالمتعلقة بالإنجليزية الأمريكية من مدونة كامبريدج الدولية Cambridge International Corpus، وهي قاعدة بيانات تحوي أكثر من مليار كلمة مكتوبة ومنطوقة." (مقدمة المعجم المذكور)

٩-٢-١-٢- - دراسات المصطلحية (علم المصطلح terminology):

التعرف على المصطلحات في المدونة، ومعرفة شيوعها في النصوص المختلفة. (انظر Bowker, Lynne and Pearson, Jennifer (٢٠٠٢)، خاصة الفصلين ٨ و ٩، وانظركذلك الدراسة التي قدمها عبدالمحسن الثبيتي في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية**، وكذلك الدراسة المقدمة من عزالدين غازي في الندوة ذاتها، وكتاب Bowker, Lynne and Pearson, Jennifer (٢٠٠٢))

ولعل أفضل كتاب شامل بالعربية يعالج دور المدونات اللغوية في دراسة المصطلحات الكتاب الذي ترجمته د. ريما بركة عن الفرنسية بعنوان: **علم المصطح: مبادئ وتقنيات** (انظر ماري-كلود لوم، ٢٠١٢، خاصة الفصول الرابع والخامس والسادس).

(انظر "الحاسوب في العمل المعجمي" في هذا الكتاب، لمزيد من المعلومات حول هذا الموضوع.)

٩-٣- اللسانية العامة:

٩-٣-١- الدراسات الصوتية:

لعل أفضل مثال لهذا النوع من الأعمال المبنية على المدونات اللغوية ماقام به رائد هذا النوع من الدراسات في هذا المجال، الدكتور علي حلمي موسى، حيث يتحدث عن العلاقة بين الصوامت والصوئت في العربية من الوجهة الإحصائية.يقول في معرض كلامه:

"والدراسة التي نعرضها اليوم تمت على عينتين من القرآن الكريم إحداهما مكية وهي سورة الأعراف وبعض قصار السور والأخرى مدنية وهي سورة البقرة. ونظراً لأن الدراسة تتم على القرآن كما نسمعه وليس كما نكتبه فقد اتبعت الطريقة الصوتية...، حيث نتبع طريقة قراءة حفص مع الوقوف على رؤوس الآيات." (انظرعلي حلمي موسى (٢٠٠١).

ومن المشروعات المهمة في مجال الأصوات المشروع الذي قام به منصور الغامدي في مدينة الملك عبد العزيز للعلوم والتقنية في الرياض. (انظر: منصور محمد الغامدي).

ويجدر بنا أن نشير إلى أن أية دراسة إحصائية عربية للجذور في المعاجم هي في واقع الأمر دراسة لأصوات العربية، خاصة الصوامت، لأن الجذر يتكون من حروف (أي أصوات). عليه فإن جميع أعمال علي حلمي موسى المتعلقة بالمعاجم العربية المعروفة تعتبر دراسات صوتية، إضافة إلى كونها معجمية.

٩-٣-٢- الدراسات النحوية والصرفية، **خاصة من منظور الاستعمال والوظائف اللغوية**:

يمكننا أن ندرس هذا الموضوع من زوايا عديدة، مثل شيوع التراكيب بعامة أو تراكيب معينة في أنواع من النصوص المختلفة. (انظر لأمثلة على ذلك: الفصلين الثالث والرابع من كتاب Biber, Conrad and Reppen, ١٩٩٨) وكذلك الفصول ١٠، ١١، ١٢، ١٦ في Aijmer and Altenberg (eds, ١٩٩١).

أمثلة للبحوث في المجالات النحوية والصرفية:

يبدو أن معظم الدراسات الأولى التي استفادت من المدونات اللغوية كانت الدراسات المعجمية. ولكن هناك توجها متزايدًا في استخدامها في الدراسات النحوية والصرفية، خاصة في بعض اللغات مثل الإنجليزية، كما نلاحظ في بايبر وآخران Biber et al. ١٩٩٨) وماير (Meyer, ٢٠٠٢) وأيجمر وألتنبيرغ (Aijmer and Alternberg, ١٩٩١). ونورد أدناه أمثلة لهذا النوع من الدراسات التي يمكن إجراؤها.

١- شيوع وتوزيع التراكيب النحوية المختلفة (الجمل الإسمية والفعلية، التركيب الوصفي والإضافي، المبني للمعلوم والمبني للمجهول، والمبني المجهول الذي يشمل الفاعل في صيغة "بواسطة" أو "مِن قِبَل" فلان أو كذا...) في اللغة بصورة عامة أوفي النصوص التي تنتمي إلى مجال معين أو منطقة جغرافية أو فترة زمنية وهكذا.

٢- ارتباط التراكيب بألفاظ أوتراكيب أخرى معينة (مثل ما نجد في العلاقة بين كان وأخواتها وصورة اسمائها وأخبارها).

٣- ارتباط بعض التراكيب، مثل أن المصدرية مقابل المصدر بأفعال معينة، أو ارتباط بعض الأفعال المتعدية إلى مفعولين بتراكيب معينة (مثلا، اصلهما مبتدأ وخبر أو أصلهما ليس مبتدأ وخبرًا).

٤- العوامل التي تؤدي إلى اختيارنا لتراكيب وتفضيلنا لها على تراكيب مشابهة في المعنى أو الوظيفة (كالجملة المبنية للمعلوم والجملة المبنية للمجهول في الأساليب الحديثة خاصة). من أمثلة ذلك كثرة ورود الجمل المبنية للمجهول في النصوص العلمية والتقنية، حيث التأكيد على العمليات والإجراءات بدلًا من الاهتمام بالفاعل لها. كذلك نجد كثرة الأفعال والصفات في النصوص القصصية، لما لهذه الجوانب من أهمية في سرد الأحداث ووصف الشخصيات في القصص.

٥- دراسة المشتقات والصيغ الصرفية المختلفة ومدى شيوعها في نصوص أو أنواع من النصوص المختلفة، وكذلك للتعرف على الأوزان والصيغ المختلفة، مثل اسم الفاعل وصيغة المبالغة "فعّال" أو المصدر الصناعي ومدى إنتاجية productivity بعضها (كثرة أو قلة استخدام وزن ما في توليد الكلمات الجديدة مثلا).

٦- دراسة السوابق واللواحق المختلفة (في الإنجليزية مثلا: in-, un-, -tion, -ity, -ic, -ism) وشيوعها في النصوص المختلفة. مثلا، أي هذه اللواصق أكثر ورودا في النصوص العلمية وأيها أكثراستعمالًا في النصوص الصحفية؟ وفي العربية، يمكننا مقارنة شيوع لواحق جمع المذكر السالم والمؤنث السالم مثلا في جمع الكلمات المقترضة (مثلًا، نلاحظ استعمال الألف والتاء في معظم الكلمات المقترضة، مثل كومبيوترات وتلفزيونات وموبايلات).

٧- دراسة شيوع أنواع الأفعال المختلفة، مثل الأفعال الدالة على الحركة (مشى،جرى) والدسالة على التغيُّر أو الثبات من جهة، والصيغ الزمنية المختلفة في أنواع النصوص المختلفة. مثلا، في البحوث والرسائل العلمية، نجد أن الباحث يستخدم صيغة المستقبل في بعض الفصول والماضي في فصول أخرى.

٨- معرفة نسبة ورود أقسام الكلمة parts of speech المختلفة في المدونة أو في أنواع النصوص المختلفة (مثلا نسبة الأفعال إلى المصادرأو الأسماء في نص أو نوع معين من النصوص).

٩- لعل من المجالات المهمة في النحو العربي دراسة قضية أقسام الكلام وتصنيف الألفاظ العربية من منظور جديد (مثلا، سبعة أقسام اقترحها تمام حسان وغيره، بدلا من الثلاثة المعروفة (انظر محمود الصيني، ١٤١٢هـ) والتثبت من صحة ذلك في ضوء المدونات اللغوية.(انظر ماذكرت أعلاه حول تصنيفيات النحويين لكثير من الأسماء والأفعال والحروف في الفقرة ١٠ من "إعداد المدونة" أعلاه.)

١٠- من الدراسات التي لم تلق العناية الكافية، وهي جديرة بالبحث في ضوء لسانيات المدونات الصيغ المختلفة التي تعبربها العربية عن الزمن والوجهة tense and aspect، مثل "كان قد فعل، قد فعل، كان يفعل، سوف يكون قد فعل..." وماشابه ذلك. (انظر١٩٨٦ Sieny, Mahmoud).

٩-٣-٣- التطور التاريخي للغة من حيث المعجم والتراكيب النحوية:

تفيدنا المدونات التايخية أو اللغوية عبر العصور في هذا النوع من الدراسات. (انظر Biber, Conrad and Reppen, ١٩٩٨: ٢٠٣-٢٢٩). كما أشرنا في مكان آخر من هذه الدراسة. من أمثلة هذه المدونات المدونة التاريخية المشهورة للغة الإنجليزية للفترة مابين ١٦٥٠ إلى ١٩٩٠، ARCHER إعداد باحثين في جامعة Northern Arizona بالولايات المتحدة الأمريكية، ومدونة Helsinki للفترة من ٨٥٠ إلى ١٧١٠ من تاريخ اللغة الإنجليزية.

ومما لاشك فيه أن أي معجم تاريخي للغة العربية لايستغني عن مدونة تاريخية شاملة وممثلة للعصور المختلفة، تقدم قدرا متوازنا من النصوص المختلفة لكل فترة زمنية (بحسب القرون أوالعصور الأدبية)، يحددها المتخصصون في اللغة والأدب العربي، مثل العصر الجاهلي، والعصر الإسلامي، والعصر الأموي....

٩-٣-٤- تحليل الخطاب discourse analysis وتحليل النصوص text analysis:

من المعروف أن هذين الفرعين من اللسانيات يتعاملان مع العلاقات فوق مستوى الجملة، بما في ذلك قضايا الترابط النصي cohesion والدلالي coherence. و من أمثلة هذا النوع من الدراسات شيوع استخدام الضمائر في النصوص الشفوية مقابل النصوص المكتوبة والمسافة بين الضمير العائد والاسم الذي يعود إليه في هذه النصوص، وكذلك اختلاف الصيغة الزمنية للأفعال في أجزاء البحوث العلمية المختلفة (المقدمة، المنهج، النتائج مثلا). (انظر: Biber, Conrad and Reppen (١٩٩٨): ١٣٢-١٠٦وكذلك (٢٠٠٨) Marcus Callies الذي يعطينا نموذجا لدراسة في هذا المجال في لغة متعلمي اللغة، و Tottie (١٩٩١). انظر أيضا محمود إسماعيل صالح (٢٠١٤)"أعمال مبنية على المدونات اللغوية العربية" لأمثلة عربية في هذا المجال.

٩-٣-٥- التداولية pragmatics و أعمال الكلام speech acts:

من الصعب أن نتصورأية دراسة لسانية تعالج الاستعمال اللغوي والعلاقة بين اللغة والوظائف والمواقف الاجتماعية في غياب مدونة مناسبة تشتمل على أمثلة واقعية لهذا الاستعمال اللغوي. ونلاحظ أنه في غياب مدونات لنصوص واقعية قد يلجأ الباحثون إلى ما يسمى بالنصوص المستنطقة elicited (مقابل التلقائية spontaneous).

من الأمثلة على هذا النوع من الدراسات البحث الذي قامت به عبير الطويل وزملاؤها في الأردن عن "المراوغة /التحاشي hedging" في الخطاب السياسي في العربية والإنجليزية. ((٢٠١١) .Taweel, Abeer et al). انظر أيضا "أعمال مبنية على المدونات اللغوية العربية" لأمثلة عربية أخرى في هذا المجال.

من النماذج من الدراسات في هذا المجال باللغة العربية مجموعة الدراسات المتعلقة بالخطاب والتداولية الواردة في كتاب: **دراسات في علم اللغة النصي: مقاربة تطبيقية على مدونة الجزيرة** (٢٠١٢).

ولمزيد من المعلومات حول المدونات اللغوية واللسانيات بصورة عامة، نقترح الرجوع إلى دراسة حديثة في هذا المجال لصالح العصيمي (٢٠١٣) بعنوان: "لسانيات المتون وعلوم اللغة".

٩-٣-٦- اكتساب اللغة وتعليمها:

٠-٣-٦-١- اكتساب اللغة الأولى والتطور اللغوي للأطفال:

كانت دراسات النمو اللغوي للأطفال تعتمد على الملاحظات الشخصية لعدد محدود من الأطفال على مدى بضع سنوات في بعض الأحيان وقد كانت محدودة من حيث العينات التي تخضع للبحث. أما مع توافر مدونات لغوية لآلاف الأطفال من أعمار مختلفة، فقد أصبح بالإمكان دراسة الموضوع على نطاق أوسع، إضافة إلى إمكانات المقارنة بين مئات الأطفال من عمر واحد أو أعمار مختلفة، مما يعطي للاستنتاجات مصداقية أعلى. كما يفيد ذلك إذا ما أجريت الدراسة على مدونات للغات مختلفة أن نتعرف على ما يسمى بالظواهراللغوية العالمية أو الكليات اللغوية language universals بصورة أدق وأفضل.

لعل من أهم المدونات اللغوية لمثل هذه الدراسة باللغة الإنجليزية مدونة "CHILDES" من إعداد قسم علم النفس في جامعة كارنيجي ميلون Carnegie Mellon; في الولايات المتحدة الأمريكية. (انظر٢٠٢-١٧٢ :(١٩٩٨).Biber et al)، حيث يورد المؤلفون أمثلة على الدراسات التي أجريت على التطور اللغوي لدى الأطفال، ويذكرون أمثلة على الملامح المميزة لكتابات التلاميذ مقارنة بالبالغين.

٩-٢-٣-٦-٢- اكتساب (تعلم) اللغة الثانية:

كنا قد أشرنا إلى أن هناك مدونات / قواعد معلومات للغات المتعلمين التي يفاد منها في هذا المجال، وذلك في الفقرة الخاصة بمدونات المتعلمين، تحت أنواع المدونات اللغوية.

(انظرأيضًا: (٢٠٠٨) John Osborne (٢٠٠٨); C Cosme)

١٠-٣-٧- تعليم اللغات:

من أبرز نماذج الإفادة من المدونات في تعليم اللغات المساهمة في اختيار محتوى المادة التعليمية (من مفردات وعبارات وتراكيب) وفي إعداد المراجع المعجمية والنحوية، كما فعل الناشر كولنز Collins في مجموعة المراجع المبنية على مدونة COBUILD تحت هذا المسمى. يضاف إلى ذلك الاستخدام المباشر للمدونات في تعليم مفردات اللغة وتراكيبها. بل إن هناك اتجاها حديثا في تعليم اللغات يسمى المنهج المعجمي lexical approach في تعليم اللغات، يعتمد إلى درجة كبيرة على المدونات اللغوية (انظر كتاب مذاهب وطرائق في تعليم اللغة: (٢٠١٤) Richards and Rogers).

جدير بالذكرأن هناك كتابين صدرا مؤخرا بعنوان: من المدونة إلى صف الدراسة (O'Keefe et al, ٢٠٠٧)، وأنماط نصية: الكلمات المفتاحية وتحليل المدونات فى تعليم اللغة (٢٠٠٦) M. Scott and Ch. Tribble كذلك تخصص سوزان هنستون (٢٠٠٢) ,S Hunston) الفصول ٦و٧و٨ من كتابها عن المدونات اللغوية واللسانيات التطبيقية(٢٠٠٢) ,Hunston).

١٠-٣-٨- الترجمة والتحليل التقابلي وتحليل الأخطاء:

١- دراسات الترجمة:

يمكن الإفادة من المدونات المتوازية خاصة في دراسة خصائص لغة الترجمة والمترجمين، وكذلك في دراسات المتقابلات في هذه المدونات. وجدير بالذكر أن هناك توجها للاستفادة من هذه المدونات في تطوير برامج الترجمة الآلية (كما فعلت شركة IBM المعروفة)، حيث تدرس الاحتمالات الاحصائية للترجمات المتقابلة في هذه المدونات لتطبيقها في الترجمة الآلية. من أمثلة المدونات التي تتعلق بالعربية المدونة التي طورتها جامعة الكويت لهذا الغرض E-A Parallel Corpus . (انظر Al-Ajmi, H., ٢٠٠٤)، وكذلك المدونة المتوازية لنصوص الأمم المتحدة: إنجليزي-عربي.

(انظر ١٢٧-١٢٣ :(٢٠٠) Hunston وكتاب (٢٠٠٤) Olohan)

٢- التحليل التقابلي contrastive analysis :

من الفوائد المعروفة لما يسمى المدونات المقارنة أو المتناظرة comparable corpora إجراء الدراسات التقابلية بين اللغات وبين التنوعات اللغوية. (انظرعلى سبيل المثال Gilquin,Papp and Diez-Bedman, ٢٠٠٨) ولا شك أن هذا يندرج تحته الدراسات المقارنة بين اللهجات المختلفة، إضافة إلى المقارنة بين اللغات، وبين لغة المتعلمين ولغة أهل اللغة الأصليين.

٣- تحليل الأخطاء:

لاشك أن هذا المجال قد يكون أول مجال في علم اللغة التطبيقي يستخدم النصوص التي ينتجها الدارسون، ما يسمى بمدونة المتعلمين learner corpus. وهناك أعداد متزايدة من مثل هذه المدونات لدارسي اللغات المختلفة من غير أهلها من خلفيات ومستويات وبيئات مختلفة، مما يتيح للباحث التعرف على الأخطاء الشائعة لدى الناطقين بلغة ما أو بلغات مختلفة، كذلك استنتاج التطور اللغوي لدى المتعلمين. ومن المعروف أن تحليل الأخطاء بطبعه يعتمد على أمثال هذه المدونات أساسا في عمله، كما

هو واضح من أي دراسة في تحليل الأخطاء في شتى اللغات. (انظر الفصل الثاني من الكتاب للتفاصيل حول حول هذا الموضوع.)

١٠-٣-٩- مجالات لسانية أخرى:

١- الأسلوبية:

يذكر بايبر وزميلاه أن من أمثلة الدراسات الأسلوبية الدراسات التي تجرى لتحديد هوية المؤلف (المجهولة أو المشكوك فيها) بمقارنة النص بنصوص أخرى معروفة المؤلف، وكذلك الدراسات التي تهدف إلى التعرف على الخصائص الأسلوبية لبعض النصوص (Biber et al., ١٩٩٨: ٢٢٣) (انظر أيضا (١٩٩١) D Crystal). ويجب أن لانسى فوائد المدونات في خدمة الكتّاب، لتحسين أسلوبهم في التعبير. تذكرهنستون أن بعض الباحثين "استخدموا المعلومات المستقاة من مدونة عامة في تقديم المشورة لكتاب الوثائق الرسمية الموجهة للقارىء العادي، من حيث الاستعمال الأنسب للغة" (Hunston, ٢٠٠٢: ١٣٥)

وجدير بالذكر أن هناك دورية قديمة تعنى بهذا النوع من الدراسات، تسمى الحوسبة الأدبية واللغوية Literary and Linguistic Computing، وقد بلغت سنتها السابعة والعشرين في العام (٢٠١٢).

٢- الأيديولوجية (العقائد الفكرية) والثقافة وعلاقتهما باللغة:

تناقش الباحثة Susan Hunston في القسم المعنون "دراسة الأيديولوجيا والثقافة Studying ideology and culture (pp. ١٠٩-١٢٣) من كتابها أمثلة لهذا النوع من الدراسات، حيث تقول إن "المدرسة السائدة في بحث العلاقة بين اللغة والعقائد (الأيديولوجيا) هي اللسانيات النقدية critical linguistics أو تحليل الخطاب النقدي critical discourse analysis الذي يدرس اللغة ليس بوصفها منظومة مستقلة ولكن بوصفها شيئا يتدخل intervenes في المجتمع، غالبا بترسيخ فرضيات وقيم ذلك المجتمع..." وتذكر أن أحد رواد هذا النوع من الدراسات (فاولر Fowler) ذكر أن "هناك ثلاثة جوانب للسانيات النقدية هي:

- دراسة النصوص في سياق الظروف الاجتماعية التي أنتجت فيها.

- الكشف عن الأيديولوجية المتضمنة implicitly coded وراء الكلام والأفكار المذكورة صراحة overt propositions.

- "تحدي الحس العام common sense بالتنبيه إلى أنه كان بالإمكان أن نمثل شيئا ما بطريقة مختلفة لها مضمون مختلف". (Hunston (٢٠٠٢). *Corpora in Applied Linguistics*: ١٠٩).

(انظر كتاب سوزان هنستون Hunston, Susan (٢٠٠٢) حيث تناقش الباحثة المجالات المختلفة من اللسانيات التطبيقية التي تستفيد من المدونات اللغوية. وكذلك، انظرمقال "المدونات اللغوية واللسانيات التطبيقية في: McEnery et al., ٢٠٠٦, pp. ٨٠-١٢٤).

ولعل من أهم الدراسات التي كتبت في هذا المجال عن اللغة العربية رسالة الدكتوراة لرجب جمعان الزهراني، والتي يقول عنها الباحث في ملخصها: " تدرس هذه الأطروحة التمظهرات الأيديوجية ideological representations للخطاب السلفي في السعودية من الفترة ١٩٨٠-٢٠٠٠ وتحاول أن تجيب عن االسؤال التالي: ما مدى وشكل تجانس الخطاب السلفي في السعودية في الفترة بين ١٩٨٠ و٢٠٠٠." وذلك في الأطروحة التي حصل بها على الدكتوراة مؤخرا (ديسمبر ٢٠١٣) من جامعة لانكاستر في بريطانيا. انظر(٢٠١٣) Alzahrani, Rajab Jamaan .

٣- اللسانيات الجنائية forensic linguistics:

من المجالات الطريفة في الإفادة من المدونات "اللسانيات الجنائية forensic linguistics" أي دراسة اللغة لأغراض قانونية وقضائية، مثل قضايا التزوير وتحديد هوية المجرمين. وتذكر هنستون أن هناك حزمة برمجية تسمى CopyCatch (اضبط النسخ!) يقارن نص "المشتبه بهم" بنص من إنتاج غيرهم (النص الضابط control text)، وتذكر أن "(٢٠٠٢) Finlay تحدد عددا من الملامح المهمة في تمييز النصوص المشبوهة من النصوص الضابطة" (Hunston, ٢٠٠٢: ١٣٣). وجدير بالذكر في هذا السياق ما أشار إليه اللساني التطبيقي كارل جيمس في كتابه عن أخطاء المتعلمين الأخطاء (Carl James, ١٩٩٨) من أن تحليل الأخطاء اللغوية يمكن استعماله للغرض نفسه.

ولاستعراض لأهم الأعمال العربية المبينة على المدونات، تنظر دراستنا المعنونة: "أعمال مبنية على مدونات لغوية عربية" في مدونة الدكتور محمود إسماعيل صالح، حيث نجد نماذج من الأعمال التي بنيت على كل من المدونات اليدوية والمحوسبة.

١٠-٣-١٠- دراسات ومشروعات اللسانيات الحاسوبية أو تطبيقاتها:

يلاحظ أن عددا من المدونات المختلفة كان وراءها خدمة البحث اللساني الحاسوبي، كما هو واضح من الدراسات ذات الطبيعة الإحصائية ونظرية الاحتمالات، كما في "تحليل إحصائي لدعم التعرف الآلي على الكتابة العربية" وكما هو الحال مع مدونتي CALLFRIEND, CALLHOME للهجة المصرية اللذين كان الغرض منهما تطوير برنامج للتعرف على الكلام، ومدونة الحياة (٢٠٠) Al-Hayat Corpus لأغراض الهندسة اللغوية واسترجاع المعلومات (انظر Al-Sulaiti, ٢٠١٠)، كذلك مشروع آي بي إم للترجمة الألية، كما أشرنا أعلاه. ويستطيع الباحث أن يلاحظ ذلك في كثير من دراسات اللسانيات الحاسوبية. (للأمثلة العربية في هذا المجال، ينظرنبيل على وعبد الغني أبو العزم والعناتي وجبر (٢٠٠٧) للدراسات المنشورة والتي قدمت في ندوات تعريب الحاسوب المختلفة.)

ويذكر الباحثان كيلغاريف وغرينفينستيت أن بوادر الارتباط الرسمي بين المدونات واللسانيات الحاسوبية يرجع رسميا إلى العام ١٩٩٣ ((٢٠٠٣) Kilgarriff, Adam and Grefenstette).

وقبل أن نختم هذا الفصل، أود التنبيه إلى أن خير مصدر للأمثلة المختلفة لدراسات لسانيات المدونات هو المجلة الدولية المعروفة باسم: المجلة الدولية للسانيات المدونات International Journal of Corpus Linguistics.

الخاتمة:

كما أشرنا من قبل في بداية هذا الفصل، فإن المدونات اللغوية ولسانيات المدونات اللغوية حقل جديد نسبيا في مجال العمل اللساني. وقد حاولنا في هذا الفصل أن نلقي الضوء على أهم جوانب هذا الموضوع، بدء بتعريف المدونات اللغوية ومرورًا بلمحة عن طرق الإفادة منها وأنواع المدونات اللغوية ومواصفاتها وإنشائها، ومتطلبات التعامل معها، ثم مناقشة تفصيلية لمجالات الإفادة من المدونات اللغوية في فروع البحث اللساني العام والتطبيقي والحاسوبي، مع التمثيل لكل فرع.

ومما لاشك فيه أن الموضوع واسع تتطلب مناقشته كتابًا أو أكثر، ولكن الغاية من هذا الفصل هو استعراض عام لهذا الحقل من حقول المعرفة لعله يكون تمهيدًا مناسبًا لهذا المجال من الحاسوب في خدمة اللغة العربية.

المراجع

أبو العزم، عبدالغني " اللغة العربية والمعالجة الآلية: برامج صخر نموذجا") في الموقع: www.aljabriabed.net/n٣١_٠٤abualazm.(٢).htm

أبو العزم، عبدالغني. الحاسوب والصناعة المعجماتية، في: **مجلة اللسان العربي**، العدد ٤٦، ١٩٩٨م، صص٢٨-٣٩.

أبو الفتوح، محمد حسين: **قائمة معجمية بألفاظ القرآن الكريم ودرجات شيوعها**. بيروت: مكتبة لبنان.

بايزيد، ليلى (٢٠١٢) "الترابط النصي في المقال الموسوم (فتاة مهمة) لسليمان العودة بين الشكل والمضمون" في: **دراسات في علم اللغة النصي: مقاربة تطبيقية على مدونة صحيفة الجزيرة** (٢٠١٢) الرياض: جامعة الأميرة نورة، ٢٩٥-٣٦٤.

بن يشو، جيلالي "حوسبة المعجم العربي: الواقع والآفاق" في: مجلة أرتين – المرجع الأول لطلاب الأدب. Art-En.com (تاريخ الوصول ٢٠١٣/٢/٢٢)

البواب، مروان (٢٠٠٩) أثر التقانات الحديثة في تجديد المعجم العربي (تاريخ الإضافة ٢٠٠٩/٣/١٨) www.aluka.net

البواب، مروان (٢٠١٢) محركات البحث في النصوص العربية. من بحوث المؤتمر السابع لمجمع اللغة العربية بدمشق (تاريخ الإضافة ٢٠١٢/٩/٢٣م) /www.aluka.net/ literature_language/٠/٧٧٢٨

بوقرة، نعمان (٢٠١٢) "استراتيجيات الإقناع في الخطاب الصحفي السعودي- دراسة نصية تداولية" في: **دراسات في علم اللغة النصي: مقاربة تطبيقية على مدونة صحيفة الجزيرة** (٢٠١٢) الرياض: جامعة الأميرة نورة، ١٩-٩٠.

التعابير الاصطلاحية في ضوء النظريات اللسانية الحديثة: دراسة تطبيقية على مدونة صحيفة الجزيرة. الرياض: جامعة الأميرة نورة.

الثبيتي، عبدالمحسن عبيد (٢٠٠٧) "استخدام ذخائر النصوص لاستخلاص المصطلحات المتخصصة" في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية**، صص ٣١-٣٨.

الثنيان، نوال (٢٠١٢) "الإحالة الضميرية- دراسة نحوية نصية" في: **دراسات في علم اللغة النصي: مقاربة تطبيقية على مدونة صحيفة الجزيرة** (٢٠١٢) الرياض: جامعة الأميرة نورة، ١٩٣-٢٩٣.

الجلاب، محمد فتحي (١٤٣٢هـ) الاستخلاص الآلي للمحتوى العربي على شبكة الإنترنت بين الواقع والمأمول، في: **السجل العلمي لمؤتمر المحتوى العربي**، صص ٨٥٣-٨٩.

حاج صالح، عبدالرحمن (١٩٩٩) "ورقة حول مشروع الذخيرة اللغوية " في: **مجلة اللسان العربي**، العدد ٤٨.

الحلوة، نوال (٢٠١٢) " أثرالتكرار في التماسك النصي- مقاربة معجمية تطبيقية" في: **دراسات في علم اللغة النصي: مقاربة تطبيقية على مدونة صحيفة الجزيرة** (٢٠١٢) الرياض: جامعة الأميرة نورة، ٩١-١٩٠.

حمادة، سلوى "المدونات النصية ودور اللغة العربية في التعامل معها". في موقع: www.thomala.com . تاريخ الحصول على الدراسة: ٢٠١٤/٨/١.

حمادة، سلوى (٢٠١١) نحو منهج عربي مُقتَرَح لتصميم المُدَوَّنات اللغويَّة. في موقع: www.globalarabnetwork.com

الحناش، محمد "برنامج لساني-حاسوبي للتعرف الآلي على التعابير المسكوكة في اللغة العربية" في: **مجلة التواصل اللساني، ملحق سلسلة الندوات**، المجلد ٣، سنة ١٩٩٦، ص ٨٩.

دراسات في علم اللغة النصي: مقاربة تطبيقية على مدونة صحيفة الجزيرة (٢٠١٢) الرياض: جامعة الأميرة نورة.

الربيعة، مها سليمان (١٤٣٣هـ) الدليل المرجعي للذخيرة النصية الفصحى لجامعة الملك سعود.
ksucorpus.ksu.edu.sa/ar/

الزركان، محمد علي (١٩٩٣) "اللسانيات و برمجة اللغة العربية في الحاسوب"، في: **وقائع المؤتمر الدولي الأول حول اللغة العربية والتقنيات المتقدمة**، الرياض، ١٠-١٤/٥/١٩٩٢، الرياض: مكتبة الملك عبدالعزيز، ١٩٩٣م، صص ٥٥-.

سالم، ثناء محمد (٢٠١٢) "الأبعاد التداولية للتعبيرالاصطلاحي" في: التعابير الاصطلاحية في ضوء النظريات اللسانية الحديثة: دراسة تطبيقية على مدونة صحيفة الجزيرة، صص ١٧-١٢٨، الرياض: جامعة الأميرة نورة.

السجل العلمي: مؤتمر المحتوى العربي في الانترنت: التحديات والطموحات (٥-٧/١١/١٤٣٢هـ) جامعة الإمام محمد بن سعود الإسلامية، الرياض. المجلد الثاني. الرياض: جامعة الإمام محمد بن سعود الاسلامية.

السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات المنعقدة في مكتبة الملك عبدالعزز بالرياض (٨-١٢ /١١/ ١٤١٢هـ). الرياض: مكتبة الملك عبدالعزيز.

السجل العلمي لندوة تقنية المعلومات والعلوم الشرعية والعربية المنعقدة في جامعة الإمام محمد بن سعود الاسلامية (١٦-١٤٢٨/٢/١٧هـ) الرياض: جامعة الإمام محمد بن سعود الاسلامية.

السلمان، عبدالملك ومنصور الغامدي وحسن الصبي (١٤٢٨هـ) نظام حاسوبي لرومنة الأسماء العربية. في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية**، صص ٢١٥-٢٢٨.

صالح، محمود إسماعيل (٢٠١٢) **الحاسوب في البحث اللغوي: لسانيات المدونات أنموذجا**. الرياض: كرسي الجزيرة للغة العربية، جامعة الأميرة نورة.

صوالحة، مجدي "العنونة الصرفية والنحوية" في الموقع التالي: www.com.leeds.ac.uk/sawalha/tagset.

صوالحة، مجدي وإرك أتول (٢٠١١) "التحليل الصرفي لنصوص العربية الحديثة والكلاسيكية"
http://www.comp.leeds.ac.uk/sawalha/sawalha١١ICCA.pdf

الصيني، محمود إسماعيل(١٤١٢هـ). نحو معجم عربي للتطبيقات الحاسوبية. في: **السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات المنعقدة في مكتبة الملك عبدالعزيز بالرياض** (٨-١٢ /١١/ ١٤١٢هـ)، صص ٥١١-٥٢١.

العارف، عبدالرحمن بن حسن "توظيف اللسانيات الحاسوبية في خدمة الدراسات اللغوية العربية: جهود ونتائج" في: **مجلة مجمع اللغة العربية الأردني**، العدد ٧٣.

عاليم، محمد (١٤٢٨هـ) المعجم العربي في ضوء اللسانيات الحاسوبية. في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية**، صص ١٥٧-١٦٦.

عبد العزيز، محمد حسن، محمد يونس الحملاوي والمعتز بالله السعيد طه (٢٠٠٨) المعجم الحاسوبي للغة العربية، بحث مقدم في: الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية المنعقد بمدينة الملك عبدالعزيز للعلوم والتقنية في الرياض (أبريل ٢٠٠٨).

عبده، داود عطية (١٤٣١هـ) في اللغة والحاسوب: الترجمة وتدقيق الإملاء بين الإنسان والآلة. عمان: دار جرير للنشر والتوزيع.

عبده، داود عطية (١٣٩٩هـ) **المفردات الشائعة في اللغة العربية**، الرياض: جامعة الملك سعود،.

العصيمي، صالح فهد (٢٠١٣) لسانيات المتون وعلوم اللغة، في: **مجلة كلية الآداب والعلوم الإنسانية** (فاس، المغرب) العدد ١٩، السنة الخامسة والثلاثون، صص ٣٧-٦٧.

علي، نبيل (١٩٨٧) اللغة العربية والحاسوب: **مجلة عالم الفكر**، م ١٨، ع ٣، ص ٧٢ـــ

علي، نبيل (١٩٨٨) **اللغة العربية والحاسوب**. الكويت: مؤسسة تعريب.

علي، نبيل (١٩٩٤) **العرب وعصر المعلومات، سلسلة عالم المعرفة**، العدد ١٨٤، المجلس الوطني للثقافة والفنون والآداب، الكويت. (ينظر الفصل التاسع من الكتاب.)

عمر، أحمد مختار (١٩٨٨) البحث اللغوي عند العرب مع دراسة في التأثير والتأثر، ط٦. القاهرة: عالم الكتب.

العناتي، وليد أحمد "اللسانيات الحاسوبية العربية: رؤية ثقافية" الموقع:
www.aljabriabed.net/n٨٢_٠٣alinati.(١).htm

العناتي، وليد وخالد الجبر (٢٠٠٧) **دليل الباحث إلى اللسانيات الحاسوبية العربية**. عمان: دار جرير للنشر والتوزيع.

غازي، عز الدين (١٤٢٨هـ) قواعد المعطيات المعرفية للمصطلحية العربية: مشروع مقترح. في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية**، صص ٢٣-٣٠.

الغامدي، عبدالله شرف و بدرية سليمان الفرهود (١٤٢٨هـ) أداة ويب معتمدة على عملية التحليل الهرمي للحصول على معجم عربي موحد لتقنية المعلومات. في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية**، صص ٣٩-٥٠.

الغامدي، منصور محمد "قاعدة بيانات الصوتيات العربية وقراءة الشفاه" www٢٦.brinkster.com/mghamdi

فايد، وفاء كامل (٢٠٠٠) "بعض مظاهر تغير الصيغ الصرفية في العربية المعاصرة" في: **بحوث الندوة الدولية للمعاجم اللغوية العامة والمختصة.(المنعقدة في الكويت في ١٤-١٧ مارس، ١٩٩٩)**

فايد، وفاء كامل (٢٠٠٧) المتطلبات اللغوية لمعالجة التعابير الاصطلاحية العربية معالجة آلية، في: **الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية**، صص ١١-٢٢.

فايد، وفاء كامل (٢٠٠٧) **معجم التعابير الاصطلاحية في العربية المعاصرة**. القاهرة: ؟؟

الفلاج، نوال (٢٠١٢) "ظاهرة الاقتران الدلالي – دراسة معجمية تطبيقية" في: **التعابير الاصطلاحية في ضوء النظريات اللسانية الحديثة: دراسة تطبيقية على مدونة صحيفة الجزيرة**، صص ١٧٧-٢٧١.

الفيفي، عبدالله يحي (٢٠١٢) المدونات اللغوية لمتعلمي اللغة العربية: نظام لتصنيف و ترميز الأخطاء اللغوية. في: **السجل العلمبي للمؤتمر الدولي لعلوم وهندسة الحاسوب في اللغة العربية في دورته الثامنة** (٢٦-٢٨ ديسمبر، ٢٠١٢) جامعة القاهرة.

القاسمي، علي محمد (٢٠٠٦). لساننيات المدونة الحاسوبية وصناعة المعجم العربي. بحث قدم إلى المؤتمر السنوي الخامس لمجمع اللغة العربية بدمشق، ٢٠-٢٢ تشرين الثاني، ٢٠٠٦.

القحطاني، سعد بن هادي. "تحليل اللغة العربية بواسطة الحاسب الآلي" **مجلة علوم اللغة**، م ٥، ع ٣، ص ٢٢٨- (القاهرة: دار غريب.)

الكمار، رأفت (٢٠٠٧) الحاسوب وميكنة اللغة العربية. القاهرة: دار الكتب العلمية للنشر والتوزيع.

لوم، ماري-كلود (٢٠١٢) **علم المصطلح: مبادئ وتقنيات، ترجمة د. ريما بركة**. بيروت: المنظمة العربية للترجمة.

المالكي، رائدة وهند القحطاني (٢٠١٢) "العلاقات الدلالية للتعبيرات الاصطلاحية في مجال الرياضة" في: التعابير الاصطلاحية في ضوء النظريات اللسانية الحديثة: دراسة تطبيقية على مدونة صحيفة الجزيرة، صص ١٢٩-١٧٧.

مراياتي، محمد، يحي مير علم، محمد حسن طيان (١٩٩٦) **المعجم الحاسوبي: إحصاء الأفعال العربية في المعجم الحاسوبي**. بيروت: مكتبة لبنان.

مهدوي، عمر (١٤٣٢هـ) المقاربة الحاسوبية للصرف العربي: قراءة في الحصيلة والآفاق. في: **السجل العلمي لمؤتمر المحتوى العربي**، صص٩٩٩-١٠٢٧.

موسى، علي حلمي و عبدالصبور شاهين.(١٩٧٣) **دراسة إحصائية لجذور تاج العروس. باستخدام الكمبيوتر**. الكويت: جامعة الكويت.

موسى، علي حلمي (١٩٧١) **دراسة إحصائية لجذور معجم الصحاح باستخدام الكمبيوتر.** الكويت: جامعة الكويت.

موسى، علي حلمي (١٩٧٢). **دراسة إحصائية لجذور لسان العرب باستخدام الكمبيوتر**. الكويت: جامعة الكويت.

موسى، علي حلمي (٢٠٠١) "حوسبة التراث العربي" محاضرة ألقيت في مجمع اللغة العربي الأردني في ١٧ نيسان (أبريل)، ٢٠٠١. www.majma.org.jo/majma/index.php/٢٠٠٩-٠٢-١٠-٠٩-٣٥-٢٨/٢٦٠-١٩-١.html/

موسى، علي حلمي (٢٠٠١) **ألفاظ القرآن الكريم: دراسة علمية تكنولوجية**. القاهرة: الشركة المتحدة للطباعة والنشر والتوزيع.

ندوة "استخدام الحاسوب في العلوم الشرعية" تحت إشراف مجمع الفقه الإسلامي (٢٤-٢٦ ربيع الآخر، ١٤١١هـ). جدة: البنك الإسلامي للتنمية.

الندوة الدولية الأولى عن الحاسب واللغة العربية: الأوراق البحثية (٢٠٠٧). الرياض: مدينة الملك عبدالعزيز للعلوم والتقنية.

النشرتي، مؤمن (١٤٣٢هـ) التحديات التي تواجه محركات البحث في استرجاع المحتوى العربي على الإنترنت: دراسة تحليلية، في: السجل العلمي لمؤتمر المحتوى العربي، صص ٧١٥-٨١٤.

هلال، يحي (١٩٩٠) الحاسوب في خدمة الحديث النبوي الشريف. في: ندوة "استخدام الحاسوب في العلوم الشرعية"، صص ٣٢٩-٣٤٢.

هلال، يحي. مختبر المعلوماتية والعلاج الآلي للغة العربية. عرض لنشاط المدرسة المحمدية للمهندسين في الرباط.

هليل، محمد محمد حلمي (٢٠٠٨)"نحو معجم عربي معاصر" من بحوث الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية. الرياض: مدينة الملك عبدالعزيز للعلوم والتقنية.

هليل، محمد محمد حلمي وسعد مصلوح وحشان العجمي (٢٠٠٠، تحرير) **بحوث الندوة الدولية للمعاجم اللغوية العامة والمختصة**.(المنعقدة في الكويت في ١٤-١٧ مارس، ١٩٩٩) الكويت: جامعة الكويت.

مراجع أجنبية:

Abdou, Ashraf (٢٠١١) *Arabic Idioms: A Corpus Based Study*. London and New York: Routledge.

Aijmer, K, and Altenberg, B (١٩٩١, eds.). *English Corpus Linguistics*. London and New York: Longman.

Al-ajmi, Hashan. (٢٠٠٤) A new English-Arabic parallel text corpus for lexicographic applications. Lexicos ١٤ (AFRILEX-reeks/series: ١٤: ٢٠٠٤, ٣٢٦-٣٣٠.

Al-Ansary, S. (٢٠٠٣). NP-Structure Types in Spoken and Written Modern Standard Arabic (MSA) Corpora. In D. Parkinson & S. Farwaneh (Eds.), *Perspectives on Arabic Linguistics XV: Papers from the Fifteenth Annual Symposium on Arabic Linguistics* (pp. ١٤٩–١٨٠). Salt Lake City.

Al-Ansary, Sameh, Nagi, Magdy and Adly, Noha Building an International Corpus of Arabic (ICA): Progress of Compilation Stage at: www.bibalex.org

Aliwy, Ahmad H. (٢٠١٣) Arabic Mophosyntactic Raw Text Part of Speech Tagging. Unpublished Ph.D. dissertation, University of Warsaw.

Al-Muhanna, Amin. (٢٠٠٤). Scientific and technological terms transfer into Arabic: A corpus-based study of Arabic noun+noun and noun+adjective compounds. Ph. D. thesis, UMIST, Manchester.

Al-Saif, Amal and Katja Markert: The Leeds Arabic Discourse Treebank: Annotating Discourse Connectives in Arabic at: www.comp.leeds.ac.uk/markert/Papers/LREC٢٠١٠-LADTB.pdf

Al-Sulaiti, Latifa and Atwell, Eric. (٢٠٠٦). "The design of a corpus of contemporary Arabic." *International Journal of Corpus Linguistics*., vol. ١١, pp. ١٣٥-١٧١.

Al-Sulaitie, Latifa(٢٠١٠). Arabic Corpora. At: http://www.comp.leeds.ac.uk./latifa/arabic_corpora.htm

Alzahrani, Rajab Jamaan (٢٠١٣). "A Corpus-Based Critical Discourse Analysis of the Ideological Representations and Legitimation in the Salafi Discourse in Saudi Arabia (١٩٨٠-٢٠٠٠)" A Ph. D. dissertation. Lancaster University, U.K.

Atkins, Sue, Jeremy Clear and Nicholas Ostler. (١٩٩٣) Corpus Design Criteria, in: *Literary and Linguistic Computing*, ٧ (١), Pp.١-١٦.

Atwell, E. (٢٠١٨). Classical and Modern Arabic Corpora: genre and language change.. In R. J. Whitt (Ed.*), Diachronic corpora, genre, and language change*.. John Benjamins.

Biber, Douglas, Conrad, Susan and Reppen, Randy (١٩٩٨). *Corpus Linguistics: Investigating language structure and use*. Cambridge: CUP.

Bowker, Lynne and Pearson, Jennifer (٢٠٠٢). *Working with Specialized language: A practical guide to using corpora*. London & New York: Routlege.

Buckwalter, Tim and Parkison, Dilworth (٢٠١١). *A Frequency Dictionary of Arabic: Core Vocabulary for Learners*. New Yok and London: Routledge, Taylor, Francis Group

Callies, M. (٢٠٠٨) "Easy to understand but difficult to use? Raising constructions and information packaging in the advanced learner variety" in: Gilquin, G, Papp, S. and Diez-Bedmar, M. (٢٠٠٨.), pp. ٢٠١-٢٢٦.

Cambridge Academic Content Dictionary (٢٠٠٩). Cambridge: Cambridge University Press.

Collins COBUILD English Language Dictionary (2006) 5th Edition. U.K.: Harper Collins Publishers.

Cosme, Ch. (٢٠٠٨) "Participle clauses in learner Englihs" in: Gilquin, G, Papp, S. and Diez-Bedmar, M. (٢٠٠٨.), pp.

Crystal, David (٢٠٠١) *Language and the Internet*, Cambridge: Cambridge University Press.

Crystal, David. (١٩٩١). "Stylistic profiling" In Aijmer and Altenberg (١٩٩١), pp. ٢٢١-٢٣٨.

Dukes, Kais and Nizar Habash "Morphological Annotation of Quranic Arabic" at: http://www.kaisdukes.com/papers/qmorph-lrec٢٠١٠.pdf

Dukes, Kais, Eric Atwell and Abdul-Baqee Sharif, "Syntactic Annotation Guidelines for the Quranic Arabic Dependency Treebank" at: http://www.comp.leeds.ac.uk/scams/papers/qsyntax-lrec٢٠١٠.pdf

Elewa, Abdul-Hamid (٢٠٠٤) "Collocation and Synonymy in Classical Arabic: A Corpus-Based Study" A Ph. D. thesis, University of Manchester.

English-Arabic Parallel Corpus of the United Nations Texts. En.wikipedia.org/…/Engish-Arabic_ Parallel_Corpus_of_United_Nations_Texts.

Facchenetti, Roberta (٢٠٠٧) *Corpus Linguistics 25 Years on*. (Language and Computers ٦٢) (Language & Computers: Studies in Practical Linguistics). Eiditions Rodopi

Francis, N.W. and Kucera, H. (١٩٧٩). *Brown Corpus Manual: Manual of Information to accompany A Standard Corpus of Present Edited American English for use with Digital Computers, Revised and Amplified*. Available at: http://icame.uib./brown/bcm.html#bc٢

Garside, R., G. Leech and G. Sampson (eds., ١٩٨٧). *The Computational Analysis of English: A corpus-based approach*. London: Longman.

Ghazali, S. & Braham, A. (٢٠٠١). Dictionary Definitions and Corpus-Based Evidence in Modern Standard Arabic. Arabic NLP Workshop at ACL/EACL,Toulouse,France. (http://www.elsnet.org/arabic٢٠٠١/ghazali.pdf)

Gilquin, G, Papp, S. and Diez-Bedmar, M. (٢٠٠٨.). *Linking up contrastive and learner corpus research*. Amsterdam and New York: Rodopi.

Granger, Sylviane (ed.) (١٩٩٨) *Learner English on Computer*. London & New York: Addison-Wesley-Longman.

Habash, Nizar, Reem Faraj and Ryan Roth. "Syntactic Annotation in the Columbia Arabic Treebank" at http://www.elda.org/medar-conference/pdf/٢٥.pdf

Harley, Trevor (٢٠٠٨) *The Psychology of Language: From data to theory, third editio*n. Hove and New York: Psychology Press (Taylor & Francis Group).

Hassan, Haslina and Nuraihan Mat Daud (٢٠١١). Corpus Analysis of Conjunctions: Arabic Learners' Difficulties with Collocations. In: *Workshop on Arabic Corpus Linguistics (WACL)*, ١١th -١٢th April ٢٠١١, Lancaster University, UK. At: http://ucrel.lancs.ac.uk/wacl/slides-HASSAN-DAUD.pdf

Holt, R. (٢٠٠٤*) Dialogue on the Internet: Language Civic Identity and computer mediated communication*. London: Praeger

Hunston, Susan (٢٠٠٢) *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

James, C. (١٩٩٨). *Errors in Language Learning and Use: Exploring Error Analysis*. London: Longman.

Kamal, Eman (٢٠٠٨) The Structure of Arguments in English and Arabic Newspaper Editorials: A Contrastive Study." An unpublished Ph. D. dissertation, King Saud University, Riyadh.

Kennedy, G.D. (١٩٩٨) *An Introduction to Corpus Linguistics*. London: Longman.

Kilgarriff, Adam and Grefenstette (٢٠٠٣)"Web as Corpus" in: *Computational Linguistics, Vol. 29*. at: www.kilgariff.co.uk./Publications/٢٠٠٣-KilgGrefesnstette-WACIntro.pdf

Lawler, J. and H.A. Dry (eds. ١٩٩٨). *Using Computers in Linguistics: A Practical Guide*. London and New York: Routledge.

Leech, Geoffrey (٢٠٠٤) Adding Linguistic Annotation in: Wynne (٢٠٠٥).

Longman Dictionary of Contemporary English (٢٠٠٩), ٥th Edition. U.K.: Pearson Longman Publishers.

McEnery, T. & Wilson, A (٢٠٠١) *English Corpus Linguistics, 2nd Ed*. Edinburgh: Edinburgh University Press.

McEnry, Tony, Xiao, Richard and Tono, Yukio (٢٠٠٦). *Corpus-Based Language Studies: An advanced resource book.* London & New York: Routledge.

Meyer, C (٢٠٠٢) *English Corpus Linguistics: An Introduction*. Cambridge: Cambridge University Press.

MOSES Machine Translation System. At http://www.statmt.org/moses

Myers, G. (٢٠١٠) *Language of Blogs and Wikis*. London: Continuum International Publishing Group.

O'Keeffe, Anne, McCarthy, Michael and Carter, Ronald (٢٠٠٧). *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: CUP.

Olohan, Maeve (٢٠٠٤). *Introducing Corpora in Translation Studies*. London: Routledge

Osborne, John (٢٠٠٨), "Adverb placement in post-intermediate English: a contrastive study of learner corpora" in Gilquin, G, Papp, S. and Diez-Bedmar, M. (٢٠٠٨.), ١٢٧٠١٤٦.

Parkinson, D. (١٩٨٥). *Constructing the social context of communication: terms of address in Egyptian Arabic*. Berlin/New York/Amsterdam: Mouton de Gruyter.

Parkinson, D. (٢٠٠٣). Future Variability: A Corpus Study of Arabic Future Particles. In D. Parkinson & S. Farwaneh (Eds.*), Perspectives on Arabic Linguistics XV: Papers from the Fifteenth Annual Symposium on Arabic Linguistics*. *Salt Lake City*, Amsterdam: John Benjamins Publishers.

Parkinson, D.& Farwaneh, S. (Eds.) (٢٠٠٣). *Perspectives on Arabic Linguistics XV: Papers from the Fifteenth Annual Symposium on Arabic Linguistics, Salt Lake City*.. Amsterdam: John Benjamins Publishers.

Richards, Jack and Theodore Rodgers (٢٠١٤). *Approaches and Methods in Language Teaching, Third Edition*. Cambridge: Cambridge University Press.

Sawalha, M. and Atwell, E. (٢٠٠٨). "Comparative Evaluation of Arabic Language Analysers and Stemmers" in: *Coling 2008: Posters and Demonstrations*, pp. ١٠٧-١١٠.

Scholfied, P. (١٩٩٥). *Quantifying Language: A Researcher's and Teacher's Guide in Computing Language Data and Reducing it to Figures*. Clevedon, U.K.: Multilingual Matters.

Scott, M. and Ch. Tribble. (٢٠٠٦) *Textual Patterns: Key words and corpus analysis in language education*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Second Workshop on Arabic Corpus Linguistics at Lancaster University (July, ٢٠١٣) See: http://www.comp.leeds.ac.uk/eric/wacl/wacl٢proceedings.pdf

Sieny, Mahmoud (١٩٨٦) "Tense and Aspect in English and Arabic: Communicative and Functional Equivalence" in: *Bulletin of the College of Arts, King Saud University*, Vol. ١٣, ٤١-٥٩.

Sinclair, J. (ed., ١٩٨٧). *Looking Up: An account of the COBUILD Project in lexical computing*. London and Glasgow: Collins ELT.

Sinclair, J. (١٩٩١) *Corpus, Concordance and Collocation*. Oxford: Oxford University Press

Smith, M. (٢٠٠٩) *Online Communication: Linking Technology, Identity and Culture*. London: Routledge.

Stubbs, M. (١٩٩٦) *Text and Corpus Analysis: Computer Assisted Studies of Language and Culture*. Oxford: Blackwell.

Taweel, Abeer Q., Saidat, Emad M Rafayah, Hussein A., & Saidat, Ahmad M.(٢٠١١). Hedging in Political Discourse, in: *The Linguistics Journal*. June ٢٠١١ Volume ٥ Issue ١, ١٦٩-١٩٦.

Taylor, S. (٢٠٠٣). Comparing Frequencies of Lexical Productions in Arabic Words. In D. Parkinson & S. Farwaneh (Eds.) *Perspectives on Arabic Linguistics XV: Papers from the Fifteenth Annual Symposium on Arabic Linguistics* (pp. ١٨١–١٨٩). Salt Lake City.

Thomson, P. (٢٠٠٤) Spoken language corpora. In: Wynne (٢٠٠٥).

Thorndike, Edward. L. and Lorge, Irving (١٩٧٢). *The Teacher's Word Book of 30,000 Word*. New York: Columbia University, Teachers College Press.

van Mol, Mark (٢٠٠٠a). The development of a new learner's dictionary for Modern Standard Arabic: the linguistic corpus approach. In U. Heid, S. Evert, E. Lehmann & C. Rohrer (Eds.), *Proceedings of the ninth EURALEX International Congress (pp. 831–836). Stuttgart, 8–12 August*. (http://www.ilt.kuleuven.ac.be/ilt/arabic/_pdf/stuttgart.pdf)

van Mol, Mark (٢٠٠٠a). The development of a new learner's dictionary for Modern Standard Arabic: the linguistic corpus approach. In U. Heid, S. Evert, E. Lehmann & C. Rohrer (Eds.), *Proceedings of the ninth EURALEX International Congress (pp. 831–836).Stuttgart,8–12August*. (http://www.ilt.kuleuven.ac.be/ilt/arabic/_pdf/stuttgart.pdf)

van Mol, Mark (٢٠٠٠b). Exploring annotated Arabic corpora: preliminary results. (http://www. ilt.kuleuven.ac.be/ilt/arabic/_pdf/tunis.pdf)

van Mol, Mark (٢٠٠٣a). Evolution of MSA, the Case of Some Complementary Particles. In D. Parkinson& S. Farwaneh (Eds.), *Perspectives on Arabic Linguistics XV: Papers from the Fifteenth Annual Symposium on Arabic Linguistics* (pp. ١٣٥–١٤٧). Salt Lake City.

van Mol, Mark (٢٠٠٣b). *Variation in Modern Standard Arabic in radio news broadcasts, a synchronic descriptive investigation into the use of complementary particles*. Belgium: Peeters.

van Mol, M.& Paulussen, H. (٢٠٠١). AraLat: a relational database for the development of bilingual Arabic dictionaries. In S. Lee (Ed.), *Proceedings of Asialex 2001, Asian Bilingualism and the Dictionary* (pp. ٢٠٦–٢١١). Seoul, August ٢٠٠١. (http://www.ilt.kuleuven.ac.be/ilt/arabic/_pdf/asialex.pdf)

West, Michael (١٩٥٣). *A General Service List of English Words: with Semantic Frequencies and a Supplementary word-list for the Writing of Popular Science and Technology*, ١١th Edition. London: Longman, Green.

Whitelock, P., M.M. Wood, H.L. Somers, R. Johnson and P. Bennett (eds., ١٩٨٧). *Linguistic Theory and Computer Applications*. London and New York: Academic Press.

Whitney, Paul (١٩٩٨.) *The Psychology of Language*. Boston and New York: Houghton Miflin Company.

Wynne, M. (٢٠٠٥) (Ed.) *Developing Linguistic Corpora: A Guide to Good Practice*. (AHDS Guides to Good Practice) At: http://www.ahds.ac.uk/creating/guides/linguistic-corpora//index.htm

لمزيد من المعلومات نحيل القارئ إلى الدراسات التالية:

سلوى حمادة "المدونات النصية ودور اللغة العربية في التعامل معها". في موقع: www.thomala.com تاريخ الحصول على الدراسة: ٢٠٠١٤/٨/١.

سلوى حمادة (٢٠١١) نحو منهج عربي مُقتَرَح لتصميم المُدَوَّنات اللغويَّة. في موقع: www.globalarabnetwork.com

صالح فهد العصيمي (٢٠١٣) لسانيات المتون وعلوم اللغة، في: **مجلة كلية الآداب والعلوم الإنسانية** (فاس، المغرب) العدد ١٩، السنة الخامسة والثلاثون، صص ٣٧-٦٧.

عبدالرحمن حاج صالح (١٩٩٩) "ورقة حول مشروع الذخيرة اللغوية " في: **مجلة اللسان العربي**، العدد ٤٨.

محمود إسماعيل صالح (٢٠١٢) **الحاسوب في البحث اللغوي: لسانيات المدونات اللغوية نموذجاً**. الرياض: كرسي الجزيرة للغة العربية، جامعة الأميرة نورة.

محمود إسماعيل صالح (٢٠١٤) "أعمال مبنية على مدونات لغوية عربية" في: مدونة الدكتورمحمود إسماعيل صالح على الشابكة.

محمود إسماعيل صالح (٢٠١٤) "لسانيات المدونات اللغوية: مقدمة للقارئ العربي" في: مدونة الدكتورمحمود إسماعيل صالح على الشابكة.

محمود إسماعيل صالح (٢٠١٤) "مصطلحات لسانيات المدونات اللغوية" في: مدونة الدكتورمحمود إسماعيل صالح على الشابكة.

ألملحق (أ)

قائمة بمدونات لغوية عربية

أولا: (ترجمة لقائمة لطيفة السليطي المؤرخة في ٢٠ فبراير، ٢٠١٠)

اسم المدونة	إنتاج	نوع اللغة	الحجم	الغرض	مصدر المادة النصية
مدونة باك – ولترالعربية ١٩٦٨–٢٠٠٣	تم ولتر	نصوص مكتوبة	٣– ٢,٥ بليون كلمة	معجمي	المصادر المنشورة على الشابكة
مُدَوَّنة ليوفان	الجـامعة الكاثوليكية في بلجيكا	نصوص مكتوبة ومنطوقة	٣ مليون منها ٧٠٠,٠٠٠ منطوقة	لعمل قاموس هولندي عربي والعكس لمتعلمي اللغة	مصادر من الشابكة، الإذاعة والتلفاز، وكتب دراسة للمرحلة الإبتدائية
مدونة نيوزواير العربية ١٩٩٤	جامعة بنسلفانيا	مكتوبة	٨٠ مليون كلمة	في التعليم وتطوير تقنياته	وكالة فرنسا للصحف ووكالة شينخوا و أمة بريس
مدونة كول فرند ١٩٩٥	جامعة بنسلفانيا	محادثات	٦٠ محادثة هاتفية	تطوير تقنيات اللغة	ناطقون باللهجة المصرية

مدونة نيميجن ١٩٦٩		مكتوبة	أكثر من ٢ مليون كلمة	العمل معجم هولندي عربي والعكس لمتعلمي اللغة	مجلات وروايات
كول هوم ١٩٩٧	جامعة بنسلفانيا	محادثات	١٢٠ محادثة هاتفية	التعرف على اللغة من خطوط الهاتف	ناطقون باللهجة المصرية
كلارا ١٩٩٧	جامعة تشالرز براغ	مكتوبة	٥٠ مليون كلمة	للعمل المعجمي	الدوريات العلمية والكتب ومصادر من الإنترنت من ١٩٩٥ حتى تاريخه
مصر ١٩٩٩	جامعة جون هوبكينز (واشنطن)	مكتوبة	غير محدد	الترجمة الآلية	مدونة ثنائية اللغة للقرآن الكريم
اسم المدونة	**إنتاج**	**نوع اللغة**	**الحجم**	**الغرض**	**مصدر المادة النصية**
الأخبار الإذاعية وكلام ٢٠٠٠	جامعة بنسلفانيا و إل دي سي	منطوقة	١١٠ برامج إذاعية	التعرف على الكلام	إذاعة صوت أمريكا
مدونة دينار ٢٠٠٠	جامعة نيميجن بهولندا و جامعة ليون دو الفرنسية بالتعاون مع سوتيتيل (التونسية)	مكتوبة	١٠٠ مليون كلمة	أغراض معجمية ومعالجة اللغات الطبيعية عامة	غير معروفة

مدونة النهار ٢٠٠١	إي إل آر إيه ELRA	مكتوبة	١٤٠ مليون كلمة	أغراض بحثية عامة	صحيفة النهار اللبنانية
مدونة الحياة ٢٠٠٢	إي إل آر إيه ELRA	مكتوبة	١٨,٦ مليون كلمة	هندسة اللغة واسترجاع المعلومات	صحيفة الحياة اللبنانية
جيجا ورد العربية	جامعة بنسلفانيا	مكتوبة	٤٠٠ مليون كلمة	معالجة اللغات الطبيعية– استعادة البيانات– نمذجة اللغة	وكالة الأنباء الفرنسية، ووكالتا أنباء النهار والحياة، ووكالة أنباء شينخوا
مدونة متوازية عر/ إنج ٢٠٠٣	جامعة الكويت	مكتوبة	٣ مليون كلمة	تدريس الترجمة والمعجمية	منشورات المجلس الوطني الكويتي
مدونة اللغة العلمية العامة ٢٠٠٤	جامعة مانشستر	مكتوبة	١٠٦ مليون	دراسة المركبات في العربية	http://www.kisr.edu.kw/science/
اسم المدونة	**إنتاج**	**نوع اللغة**	**الحجم**	**الغرض**	**مصدر المادة النصية**

مدونة العربية الفصحى ٢٠٠٤	جامعة مانشستر	مكتوبة	٥ مليون	تحليل معجمي	www.muhaddith.org and www.alwaraq.com
مدونة متعددة اللغات ٢٠٠٤	اج مانشستر	مكتوبة	١٠٧ مليون (منها ١ مليون عربية)	الترجمة	صفحات تقنية المعلومات
مدونة سوتيتل (التونسية)	سوتيتل لتقنية المعلومات (تونس)	مكتوبة	٨ مليون كلمة	معجمي	مواد أكاديمية أدبية ومجلات علمية

المصدر : http://www.comp.leeds.ac.uk/latifa/arabic_corpora.htm

ثانياً: مدونات لغوية إضافية:

يضاف إلى المدونات المذكورة أعلاه المدونات التالية:

اسم المدونة	إنتاج	نوع اللغة	الحجم	الغرض	مصدر المادة النصية
الذخيرة النصية الفصحى لجامعة الملك سعود	مها سليمان الربيعة	نصوص مكتوبة	٥٠٦٠٢٤١٢ كلمة	أغراض مختلفة	مصادر تراثية مختلفة
مدونة "تونسية: مدونة العربية التونسية	كارين ماكنيل وفايزة ميلاد	نصوص محكية			
مدونة خليج ٢٠٠٤	مراد عباس	نصوص مكتوبة	٤,١ ميقا بايت	معالجة اللغات الطبيعية	حوالي ٢٠٠٠٠ مقالة من الخليج
المدونة العربية arabiCorpus	جامعة بريغهام يونغ (الأمريكية)	نصوص مكتوبة ومنطوقة	١٧٣,٦٠٠,٠٠٠ كلمة	أغراض مختلفة	صحافة، أدب، حديث، نثر، عامية مصرية...

المدونة اللغوية العربية	مدينة الملك عبدالعزيز للعلوم والتقنية (الرياض)	نصوص مكتوبة	٧٣٢٧٨٠٥٠٩ كلمة	لأغراض مختلفة	مصادر متنوعة قديمة وحديثة وتمثل مناطق جغرافية مختلفة
المدونة اللغوية العربية الدولية	مكتبة الإسكندرية	نصوص مكتوبة	متوقع ١٠٠ مليون كلمة	أغراض مختلفة	مصادر مختلفة
المدونةاللغوية لمتعلمي اللغة العربية	الفيفي وأتويل	نصوص مكتوبة لدارسي العربية	٢٨٢,٧٣٢ كلمة	بحث لغوي	كتابات الطلاب
مدونة متعلمي العربية	جامعة أريزونا	نصوص مكتوبة		تحليل أخطاء المتعلمين	كتابات دارسين أجانب على مدى ١٥ عاما

المدونة المتوازية (إنجليزي-عربي)	حمودة الصالحي	نصوص مكتوبة	٣٧٩٤٦٧٧ كلمة إنجليزية و ٣٧٥٥٧٤١ كلمة عربية	بحث لغوي	نصوص مترجمة في الأمم المتحدة
مدونة مسكوك (arabcrunch)	طه الزروفي (الجزائر)	نصوص مكتوبة، متلازمات لفظية		بحث لغوي	وكيبيديا العربية ومكتبة ثواب
مدونة وطن ٢٠٠٤	مراد عباس	نصوص مكتوبة		معالجة اللغات الطبيعية	٥٠٠٠ مقالة من الوطن
مكنز صخر	شركة صخر	نصوص مكتوبة	متوقع ٥٠٠ مليون كلمة	تطوير برمجيات حاسوبية لغوية	نصوص نثرية معاصرة
KALIMAT: A Multipurpose Arabic Corpus		مدونة نصوص عربية		متعددة الأغراض	

KALIMAT: A (٢) Multi-document Summaries Corpus		مدونة لملخصات وثائق عربية			
Gumar corpus	جامعة نيويوك – أبو ظبي	عام١٢٠٠ نص قصصي	١١٢ مليون كلمة	دراسة لهجات الجزيرة العربية	نصوص قصصية من الجزيرة العربية من الشابكة

المصادر:

- عبدالغني أبو العزم، اللغة العربية والمعالجة الآلية: برامج صخر نموذجاً:

www.aljabriabed.net/n٣١_٠٤abualazm.(٢).htm

- مدونتا الوطن والخليج، مرادعباس:

http://sites.google.com/site/mouradabbas٩/corporal

مدونة "تونسية: مدونة العربية التونسية للباحثتين كارين ماكنيل وفايزة ميلاد:

www.tunisiya.org

- مدونة العربية arabiCorpus:

http://arabicorpus.byu.edu

- مدونة كلمات:

www.lancs.ac.uk/staff/elhaj/corpora.htm

- مدونة متعلمي العربية:

http://l٢arabiccorpus.cercll.arizona.edu

(لمزيد من المعلومات حول مدونات متعلمي اللغة، ينظر عبدالله الفيفي (٢٠١٥))

- مدونة مسكوك، في الموقعين:

/http://arabcrunch.com/ar

http://maskouk.sourceforge.net/index.php?content=doc

- مدونة نصوص الأمم المتحدة لحمودة الصالحي انظر:

en.wikipedia.org/…/Engish-Arabic_
Parallel_Corpus_of_United_Nations_Texts

- مدينة الملك عبدالعزيز للعلوم والتقنية (الرياض):

/http://www.kacstac.org.sa

- مها سليمان الربيعة (١٤٣٣هـ) الدليل المرجعي للذخيرة النصية الفصحى لجامعة الملك سعود:

/ksucorpus.ksu.edu.sa/ar

ثالثاً: مدونات عربية موسومة:

لعل من أهم المدونات العربية الموسومة (أي بها تحشية صرفية ونحوية) مايلي:

(١) مدونة القرآن الكريم، وهي موسومة صرفياً ونحوياً وبها شبكة دلالية للمفاهيم ontology. وهي من إعداد فريق من الباحثين من جامعة ليدز (مجموعة البحث اللغوي)، وهي في نسختها الرابعة في الأول من مايو ٢٠١١. الموقع: http://corpus.quran.com

(٢) مدونة جامعة بنسلفانيا المشكولة والموسومة بأقسام الكلام والمحللة نحوياً، من إعداد باحثين من جامعة بنسلفانيا بمدينة فيلادلفيا الأمريكية برئاسة محمد المعموري، وتسمى Penn Arabic Treebank:

Arabic Treebank: Part ١ v ٣،٠ (POS with full vocalization + syntactic analysis). Authors: Mohamed Maamouri (project head), Ann Bies, Tim Buckwalter and Hubert Jin.

(٣) مدونة جامعة كولومبيا الموسومة صرفياً ونحوياً CATiB، وهي من إنتاج باحثين من جامعة كولومبيا في نيويورك. انظر Nizar Habash, Reem Faraj and Roth, Syntactic Annotation in the Columbia Arabic Treebank

وكذلك: CATiB: The Columbia Arabic Treebank (٢٠٠٩) in: Proceedings of the ICL-IJNLP ٢٠٠٩ Conference Short Parpers.

(٤) مدونة براغ الموسومة نحوياً Prague Arbic Dependency Treebank من إعداد فريق بحث في جامعة تشارلس Charles Universtiy في براغ Prague. وتشتمل المدونة على ١١٣ ألف وخمسمئة كلمة فعلية/نصية، معظم نصوصها من إصدارات وكالات للأنباء:

http://ufal.mff.cuni.cz/padt/PADT_١،٠/docs/index.html

(٥) الذخيرة النصية الفصحى الموسّمة: تذكر منشئة المدونة الباحثة مها الربيعة (في ٧ يونيو ٢٠١٤) عن المدونة اللغوية المذكورة ما يلي: "هي نسخة من الذخيرة النصية الفصحى تحتوي على وسوم توضح الترويسة المعجمية، الجذع، الجذر، الجنس والعدد لكل كلمة من كلمات الذخيرة النصية. (الدليل المرجعي للذخيرة النصية الفصحى الموسّمة: ksucorpus.ksu.edu.sa/ar/?p=١١٧).

(٦) مدونة قُمَر Gumar corpus. يذكر معدو المدونة المذكورة بأنها موسومة صرفيًا. انظر:

https://nyuad.nyu.edu/en/research/faculty-labs-and-projects/computational-approaches-to-modeling-language-lab/research/gumar-corpus.html

مدونات إضافية اخرى

Quranic Arabic Corpus

QurAna: Quran pronoun anaphoric co-reference corpus

QurSim: Quran verse similarity corpus

Qurany: Classical Arabic Quran with English translations and verse topics

Boundary-Annotated Quran Corpus

Quran Question and Answer Corpus

Multilingual Hadith Corpus

KSUCCA King Saud University Corpus of Classical Arabic

ABC: Arabic By Computer

CCA: Corpus of Contemporary Arabic

Arabic Internet Corpus

World Wide Arabic Corpus

Arabic Discourse Treebank

Arabic Learner Corpus

Arabic Children’s Corpus

Arabic Dialect Text Corpus

BBC Arabic corpus:

CNN Arabic corpus:

الملحق (ب)

مصطلحات لسانيات ا لمدونات: إنجليزي-عربي

وقائمة عربية-إنجليزية

نظرًا لحداثة لسانيات المدونات اللغوية، رأينا إضافة هذا الملحق ليستفيد منها الباحث عند قراءة المراجع الإنجليزية ويفيد منها المترجمون، ويتعرف الجميع على المقابلات العربية والإنجليزية لأهم مصطلحات هذا الفرع من العلوم اللغوية الحديثة.

أولا: المسرد الإنجليزي-العربي

محاذاة: alignment

وضع الترجمة (عادة، على مستوى الجملة أو العبارة أو الكلمة) بجوار مقابلتها في لغة أخرى، غالبا في ذاكرات الترجمة والمدونات المتوازية.

قائمة ألفبائية: alphabetical list

قائمة ألفاظ مرتبة حسب حروف الهجاء، مقابل قائمة مرتبة حسب تكرار الكلمات.

تحشية الضمير العائد: anaphoric annotation

وسم نصوص المدونة بالمعلومات الخاصة بالضمائر والأسماء التي تعود عليها.

مدونة بها تحشية: annotated corpus

المدونة التي تم وسمها بالمعلومات اللغوية (نحوية أو دلالية مثلا).

تحشية/ تعليقات: annotation

في لسانيات المدونات، يقصد بها تزويد النصوص بالمعلومات اللغوية، وتكون داخل النص.

يزود النص بالتحشية: **annotate**

الذكاء الاصطناعي: **Artificial Intelligence (AI)**

فرع من فروع علوم الحاسوب يستعمل لحل المشكلات، عن طريق محاكاة السلوك الإنساني في التعامل مع المعرفة.

مطابقة الواقع: **authenticity**

أي تكون المادة اللغوية مستقاة من نصوص حقيقية، وليست مصطنعة.

مدونة متوازنة: balanced corpus

مدونة جمعت نصوصا وفق معايير تحقق التوازن بين أنواع النصوص المختلفة من حيث النوع والجنس الأدبي والتخصص...

طقعم أساس للواسمات: base tagset

مجموعة أساسية من الرموز تستخدم في وسم النصوص. ومن أشهرها "مشروع تشفير النصوص Text Encoding Initiative (TEI).

كشاف سياقي ثنائي اللغة: bilingual concordance

كشاف سياقي يشتمل على نتائج البحث في مدونة متوازية، يعرض للمقاطع النصية مع مقابلاتها، إما جنبا إلى جنب أو فوق بعضها البعض.

مصطلح محتمل: candidate term

كلمة أو عبارة يشتبه في كونها مصطلحا عثر عليه برنامج استخراج المصطلحات.

سي إي إس: CES

اختصار لمعيار تشفير المدونات Corpus Encoding Standard. وهو تطبيق للغة إس جي إم إل SGML وينسجم بشكل عام مع سي إي إس.

محرف: character

هو الرمز الذي يشير إلى حرف أو رقم في الحاسوب

تشفير المحارف: character encoding

نظام استخدام قيم رقمية لتمثيل المحارف (الحروف أو الأرقام).

أدوات اللسانيات الحاسوبية: CL tools

البرمجيات التي تشمل التحليل الصرفي وتحديد البيئة المعجمية وما شابه ذلك.

مصاحب لفظي: collocate

الكلمات التي تسبق أوتأتي بعد كلمة معينة (مثلا: حبة خردل، لانقول مثلا قطعة خردل).

يصاحب لفظيا: collocate with

تسبق أوتأتي بعد كلمة معينة (مثلا: نجد في القرآن الكريم أن "ذرة" يصاحبها لفظ "مثقال" ، وليس "وزن" فالعبارة هي:"مثقال ذرة".).

مصاحبة لفظية: collocation

مجموعةالكلمات التي تتلازم في وقوعها في النصوص (مثل: باقة من الزهور، فريق من اللاعبين، قطيع من الماشية، سرب من الطيور). ويذكر بعضهم بأن المسافة بين الكلمتين لا تتعدى أربع كلمات عادة (مثل، قطيع كبير جدا من الماشية).

تلازم نحوي: colligation

أنماط المصاحبة اللفظية بناء على معايير نحوية وليست دلالية، مثل ورود كلمة ما مع فئة نحوية من الكلمات (أفعال أو أسماء أوحروف مثلا).

مدونة مشابهة: comparable corpus

مدونة للمقارنة بين أكثر من لغة، ويشتمل على عدد من المدونات في كل لغة تشبه في تكوينها composition pattern مدونات اللغة الأخرى، أي تستخدم فيها نفس معايير اختيار النصوص.

مدونة مقارنة: comparative corpus

مدونة تشتمل على مكونات أو نصوص من لهجات لغة واحدة.

جمع compilation:

جمع النصوص مثلا.

اللسانيات الحاسوبية: Computational Linguistics (CL)

فرع من اللسانيات تطبق فيه تقنيات ومفاهيم حاسوبية للتعامل مع المشكلات الصوتية واللغوية الأخرى.ويتطلب معرفة جيدة بالبرمجة الحاسوبية.

كشاف سياقي: concordance

قائمة تورد جميع ألفاظ أوعبارات النص في سياقاتها (أي مسبوقة بعدد من الكلمات ومتبوعة بعدد منها). وهو مايعرف بمصطلح الكلمة المفتاحية في سياق Key Word in Context (KWIC).

سياق: context

الكلمات المحيطة بالكلمة المقصودة.

بحث بالسياق: context search

نوع من البحث يتيح لنا البحث عن نمط بحث (كلمة مثلا) في جوار نمط بحث آخر (مثلا: ابحث عن كلمة "بصري" في جميع الحالات التي ترد فيها في مدى خمس كلمات من "قارىء").

مدونات: جمع مدونة corpora:

مدونة:corpus

مجموعة من النصوص اللغوية في صورة محوسبة / رقمية. ويطلق عليها البعض اسم الذخيرة، كما أن هناك حالات قليلة نجد فيها كلمة مكنز(الخاطئة).

توازن المدونة: corpus balance

اشتمال المدونة على النصوص المطلوبة بشكل عينات متوازنة.

ترويسة المدونة: corpus header

المعلومات الببليوغرافية والتصنيف وأية معلومات عن المدونة.

لسانيات المدونات (اللغوية): Corpus Linguistics

فرع من اللسانيات يعتمد في دراساته على المدونات اللغوية، أي النصوص المكتوبة والمنطوقة للعة. (انظر الشرح في صلب الدراسة.)

معالجة المدونة: corpus processing

مصطلح يشير إلى جميع العمليات التي يقوم بها الباحث اللساني بالمدونة.

نص: co-text

يقصد به أي مادة لغوية ترد فيها الكلمة أوالعبارة المعنية، أي سياق لغوي.

جدولة مقطعية:cross-tabulation

جدول يبين شيوع ظاهرة لغوية في عدد من العينات.

مدونة تاريخية: diachronic corpus

مدونة تشتمل على نصوص من عصور مختلفة، وتسمى أحيانا historical corpus.

إزالة الغموض: disambiguation

ما يفعله الباحث مثلا من عمليات تعين في إزالة اللبس أو الغموض، كما في المشتركات اللفظية أو اللبس التركيبي، ويشمل مثلا تحديد المعنى المقصود وفقا للسياق أو تحديد نوع الكلمة أو قسم الكلام الذي تنتمي إليه (فعل أو صفة أو اسم مثلا).

تحشية خطابية: discoursal annotation

تحشية تضاف في داخل النص لتحديد جوانب من ملامح الخطاب، مثل العلاقات بين الجمل أو الفقرات أو الضمائر والأسماء التي تعود عليها وغيرذلك.

وسم مطابق:ditto tag

وسم بنفس السمة:ditto tagging

وسم جميع مكونات التعبيرة الاصطلاحية بنفس السمة لبيان انتمائها إلى نفس الوحدة التعبيرية.

دي تي دي:DTD

اختصارلعبارة Document Type Definition أي تحديد نوع الوثيقة. هونوع من المعلومات المشفرة التي تخبر المستعمل أو البرنامج الحاسوب عن العناصر التي يحتويها النص وكيف تتجمع مع بعضها البعض، ويتكون من طقم أساسي من السمات، إضافة إلى أطقم من السمات الخاصة التي يختارها المستعمل التي يختارها مما يتم إعداده وفقا لمجموعة القواعد المذكورة في التعليمات العامة لمبادرة تشفير النصوص (TEI Guidelines).

تشفير: encoding

مصطلح يطلق على عملية تمثيل البيانات النصية واللغوية (أي التحشية والسمات) بصيغة معينة في المدونة، والتي يجب أن تكون وفقا لمعايير معروفة حتى يسهل إعادة الاستفادة من مجموعات النصوص وتبادلها بين الباحثين.

وسم الأخطاء::error tagging

عملية إعطاء رموز في النص تبين أنواع الأخطاء في مدونة متعلمي اللغة، للاستفادة منها عند تحليل الأخطاء في المدونة.

قائمة شيوع: frequency list

القائمة التي تورد كلمات النص حسب شيوعها أو تكرارها في النص، مع بيان عدد مرات ورودها. ويمكن ترتيب القائمة في ثلاث صور مختلفة: (١) حسب شيوع الكلمات تنازليا، أي من الأكثر شيوعا إلى الأقل شيوعا، (١) حسب شيوع الكلمات تصاعديا، أي من الأقل شيوعا إلى الأكثر شيوعا، (٣) حسب الألفاظ مرتبة ألفبائيا، مع ذكر تكرار كل كلمة بجوارها.

شيوع/ تكرار frequency:

مرات ورود ظاهرة ما (مثلا، كلمة أو تركيب أو ملمح) في النص أو المدونة.

إعراب كامل: full parsing

إعراب جمل النص بالتفصيل، أي بيان الوظيفة النحوية لكل كلمة أو تركيب فيه.

مدقق عام للنص: general text checker

برنامج يدقق الجوانب العملية في النص، مثل تعليمه بصورة مناسبة وبيان بدايات الجمل مثلا ونهاياتها.

جنس أدبي/ نوع: genre

المقصود به نوع النص، من حيث هو شفوي أومكتوب أوهو أدبي (قصصي، مسرحي، شعري) أوعلمي وتقني أوإعلامي (صحفي، إذاعي)...إلى غيرذلك. ويلاحظ أن المصطلح يستعمل في لسانيات المدونات بمعنى أكثرعمومية من "الجنس الأدبي".

ترويسة:header

مصطلح يطلق على مايرد في بداية النص، مثل اسم المؤلف والعنوان والمصدر...إلخ.

تحليل تحاوري:interactive analysis

التحليل الذي يشترك فيه الحاسوب مع الإنسان (اللساني مثلا)، حيث يقوم الحاسوب بجزء من التحليل ويقوم الإنسان بالجزء الآخرمن التحليل (مثل فك اللبس).

- **كلمة مفتاحية: .keyword**

الكلمة الهدف موضوع الفحص أو التحليل.

سياق غني بالمعلومات: Knowledge-rich context

سياق يشتمل على تعريف أوشرح لمصطلح أو مفهوم ما.

كويك (الكلمة المفتاحية في سياق Key-Word In Context**) KWIC**

الصيغة التي ترد فيها الكلمة في الكشاف السياقي، حيث ترد الكلمة المفتاحية أو الهدف في وسط السطر مسبوقة ومتبوعة بعدد من الكلمات.

هندسة لغوية: Language engineering

تهدف الهندسة اللغوية إلى تيسير استخدام التطبيقات المختلفة وزيادة احتمالات التواصل بين لغات العالم بدمج الطرق الجديدة لمعالجة اللغة الشفوية والمكتوبة، مثل إنشاء الوثائق وإدارتها وخدمات المعلومات والاتصالات والترجمة واكتساب اللغات الأجنبية.

موارد/ مصادر لغوية: Language/Linguistic Resources

مصطلح عام يطلق على الموارد أو المصادر مثل مدونات اللغة المنطوقة أو المكتوبة وقوائم الشيوع وقائمة الألفاظ والأدوات المستخدمة للحصول على المعلومات اللغوية.

مدونة المتعلمين: learner corpus: a corpus of language produced by learners مدونة تشتمل على نصوص من إنتاج دارسي اللغة من غير أهلها (غالبا).

مرتِّب حسب اليسار: left sorted

يعني في الإنجليزية، مرتب حسب الكلمة السابقة للكلمة المفتاحية. أما في العربية فيقصد به وفق الكلمة التالية لها.

كلمة أساس:lemma

الكلمة في أبسط صورها المجردة من الزوائد (مثل الاسم المفرد النكرة والفعل في صيغة الماضي للمفرد المذكر الغائب)، وهو مانجده في مداخل المعجم.

تجريد الكلمة:lemmatization

إعادة الكلمة إلى صورتها الأساس، أي حذف الزوائد ورد الكلمة المتصرفة إلى صورتها الأساس (مثل تحويل يكتبان ويكتبون واكتب... إلى الفعل "كتب" وتحويل الاسم في صيغتي المثنى والجمع إلى المفرد).

وحدة معجمية: lexeme

غالبا ما يطلق على الكلمة الأساس. ويستخدم البعض المصطلح ليشمل التعبيرة الاصطلاحية (المسكوكة) التي تؤدي معنى معجميا معينا.

قاعدة بيانات معجمية:lexical data base (LDB)

قاعدة بيانات للألفاظ مخزنة وفقا للأسس المتبعة في قواعد البيانات، من حيث اشتمالها على سجلات records وحقول fields تحتوى على بيانات تعيننا في استرجاع المعلومات.

Lexical Resources/Dataمصادر/ بيانات معجمية:

مصطلح يطلق على البيانات المعجمية المخزنة حاسوبيا والتي نستفيد منها في البحث المعجمي أو في إنتاج المعاجم.

lexicographyصناعة المعاجم/ المعجمية:

إعداد المعاجم المختلفة: ويشمل استخدام المدونات لحصر مداخل المعجم وجمع المعلومات الدلالية والإملائية والصرفية والشواهد...إلخ، وعرضها في صورة منتظمة (مرتبة ألفبائيا أوحسب الموضوعات) مع كافة وسائل الشرح اللازمة لبيان معاني المداخل واستعمالاته. ويقصد بالمصطلح نفسه العلم الذي يدرس هذا الفن.

معجم:lexicon

كلمة مرادفة لقاموس أو معجم، أي قائمة بألفاظ اللغة ومعلومات عنها، وتستعمل بدلا عنهما عندما تكون محوسبة. ويستخدم في اللسانيات النظرية ليشير إلى القاموس الذهني mental لمتحدث اللغة.

اللغة للأغراض العامة: LGP

اختصار لعبارة language for general purposes اللغة غير المتخصصة بمجال معين.

اللغة لأغراض خاصة: LSP

اختصار لعبارة language for special purposes اللغة التي تستخدم في معالجة مجالات المعرفة المتخصصة، أو التي تستخدم لأغراض خاصة مثل اللغة لأغراض أكاديمية التي يحتاج إليها مثلا الطالب الجامعي.

تعليم النص: mark-up

علامات أورموز تستخدم في النص لتقديم معلومات عنه، مثل مصدره ومؤلفه وتاريخ المصدر، وكذلك السمات المعجمية والنحوية. غير أن هناك من يرى قصر المصطلح على المعلومات خارج النص (كالمصدر والتاريخ)، مقابل تحشية annotation التي تكون في داخل النص.

بيانات عن (شيء ما) :metadata

يطلق على البيانات أو المعلومات عن شئ ما (كلمة أوجملة أو نص) ولايكون جزءا منه.

ومن أمثلتها اللغة التي نستخدمها في كتب النحو مثلا لوصف التراكيب اللغوية، وتسمى metalanguage (حرفيا: لغة ماورائية).

مدونة مراقبة (التطور): monitor corpus

مدونة يتم تحديثها سنويا لمتابعة أية مستجدات على اللغة، وتستخدم لتحديث المعاجم.

مدونة أحادية اللغة: monolingual corpus

مدونة تشتمل على نصوص من لغة واحدة.

مدونة متعددة اللغات:multilingual corpus

مدونة تشتمل على عينات من النصوص في لغات مختلفة.

معالجة اللغات الطبيعية: Natural Language Processing (NLP)

معالجة اللغة البشرية والتعامل معها حاسوبيا لأغراض مختلفة.

عنصر هدف/ عقدة:node

العنصر(كلمة أو تعبير) قيد الدراسة. وفي التحليل الشجري للجمل، يعني العقدة التي تتفرع منها التراكيب المختلفة.

تسوية: normalization

إجراء يقوم به الباحث للتسوية إحصائيا بين نصين مختلفين في الحجم عند مقارنة عدد الكلمات فيها، وذلك بذكرعدد الكلمات في كل مائة أو ألف كلمة في كل نص. فيقال مثلا أن كلمة ما ترد كذا مرة في كل مائة أو ألف كلمة في كل نص.

مدونة متوازية (محاذاة): parallel (aligned) corpus multilingual.

مدونة تشتمل على نصوص متوازية أي تمت محاذاة جملها أوعباراتها في لغتين، وغالبا مايكون أحد النصين ترجمة للأخرى. ويستفاد منها في دراسة بعض قضايا الترجمة مثل لغة الترجمة ودقة الترجمة والتحليل التقابلي والترجمة شبه الألية إلى غير ذلك.

معرِب نحوي:parser/syntactic parser

برنامج يقوم بإعراب الجمل، وذلك بالتعرف على أنواع الكلمات والوحدات التركيبية ووظائفها في الجملة.

إعراب: parsing

تحليل الجملة نحويا، كما أشرنا أعلاه.

مدونة معرَبَة:parsed corpus

مدونة تم تحليلها نحويا، أو إعراب الجمل فيها، مع وسم النص نحويا.

واسم لأقسام الكلام:part-of-speech tagger

برنامج يقوم بتعيين نوع الكلمة (قسم الكلام) آليا. فيسم كل كلمة مبينا نوعها، اسما كانت أم فعلا أم حرفا، مثلا. وهو مطلوب في كثيرمن أنواع البحث المعجمي وفي الإعراب الآلي.وتسمى العملية وسم قسم الكلام part-of-speech tagging.

كتابة صوتية:phonetic transcription

كتابة النصوص الصوتية باستخدام رموزصوتية تمثل الكلام. وقد يشمل تسجيل الملامح فوق القطعية، مثل النبر والتنغيم والوقف.

وسم موجه نحو المشكلة: problem-oriented tagging

نوع من التحشية للنص بغرض حل مشكلة أو مشكلات محددة يرغب الباحث في معالجتها.

أدوات (بحث) مشاعة:public domain tools

أدوات بحث متاحة للاستعمال العام، ولاتتطلب تصريحا بذلك.

أدوات استعلام: query(ing) tools

أدوات تستعمل للاستعلام واسترجاع البيانات والمعلومات، مثل محركات البحث search engines.

مدونة خام: raw corpus

مدونة لم تتم تحشيتها بأية سمات.

إمكانية الاستعادة:recoverability

إمكانية استعادة النص الأصل من المدونة التي تم تحشيتها annotated.

مدونةمرجعية: reference corpus

مدونة صممت بناء على معايير متفق عليها، وتشمل اللغة المحكية والمكتوبة والرسمية وغير الرسمية، تمثل مستويات اللغة المختلفة. وتعتبر معيارا مرجعيا. من ثم يحب أن لايقل عدد كلماتها عن خمسين مليون كلمة.

تمثيل (اللغة):representativeness

اشتمال المدونة على نوع اللغة التي يفترض أن تمثلها، مما يتيح تعميم الاستنتاجات المستخلصة من دراستها.

مرتِّب حسب اليمين: right sorted

يعني في الإنجليزية، أن الكشاف السياقي مرتب حسب الكلمة اللاحقة للكلمة المفتاحية. أما في العربية فيقصد به وفق الكلمة السابقة لها.

مدونة عينة:sampled corpus

مدونة محدودة، يتم اختيارها بعناية فائقة وتدرس بالتفصيل.

مدونة مشبّعة:saturated corpus

مدونة بلغت حدا تصبح زيادة الألفاظ النوعية الجديدة محدودا إليها عند إضافة أية نصوص جديدة.

محرك بحث: search engine

برنامج يعيننا في البحث عن كلمات أوعبارات معينة في المدونة أو الشابكة (الانترنت)، مثل Google.

نمط بحث: search pattern

أي عنصر لغوي نبحث عنه، سواء أكان كلمة أم عبارة.

تحشية دلالية: semantic annotation

تحشية لبيان العلاقات والملامح الدلالية (معاني الكلمة).

إحصاء دلالي: semantic count

إحصاء المعاني المختلفة للكلمات وتكرارها في النص.

المعنى الناتج:semantic prosody

المعنى الذي ينتج من التفاعل بين الكلمة ومصاحباتها اللفظية.

إس جي إم إل: SGML Markup Language

اختصار للمصطلح لغة التعليم العامة المعيارية Standard Generalized Markup Language التي تعتبر معيارا دوليا لتشفير النصوص المحفوظة إلكترونيا.

إعراب جزئي: skeleton parsing

إعراب يتجاهل التفاصيل المتبعة في الإعراب الكامل للنص، ويسمى أحيانا بالإعراب السطحي.

ترتيب: sort(ing)

تنظيم الكلمات مثلا في ترتيب معين.

مدى: span

عدد الكلمات السابقة للكلمة المفتاحية واللاحقة له (مثلا أربع كلمات سابقة وأربع كلمات لاحقة).

مدونة خاصة: special corpus

مدونة تصصم لغرض محدد، مثل نص معين، من ثم لا تعتبر ممثلة للغة بصورة عامة.

مدقق إملائي: spelling checker

برنامج يراجع الصحة الإملائية لكلمات النص.

مدونة محكية: spoken corpus

مدونة للغة المحكية، يراعى فيها عادة الكتابة الصوتية للنص.

قائمة الاستبعاد: stop list

قائمة الكلمات التي نريد الحاسوب أن يسبعدها عند إحصاء تكرار الكلمات، وهي غالبا الكلمات الوظيفية مثل حروف الجر والضمائر وسائر الأدوات (نظرا لعلمنا بشيوعها).

سلسلة::string

سلسلة أومجموعة متتابعة من الحروف أوالمحارف.

مدقق أسلوبي:style checker

برنامج لتدقيق الجوانب الأسلوبية للنص.

مدونة فرعية::subcorpus

جزء مقتطع من المدونة يتسم بخصائص معينة.

مدونة تزامنية::synchronic corpus

مدونة تشتمل على نصوص من فترة زمنية معينة، مقابل مدونة تاريخيةdiachronic corpus

سمة:tag

عنوان أوعلامة تضاف إلى كلمة ما أو أية وحدة لغوية أخرى تعطي معلومات عنها (مثلا: كتب-اسم، كتب-فعل).

وسم: tagging

عملية إضافة سمات للوحدات اللغوية تعين في عملية التحليل اللغوي للنص أوالمدونة.

طقم السمات:tagset

مجموعة السمات المستخدمة في وسم المدونة.

إرشادات التي إي آي:TEI Guidelines

الإرشادات الخاصة بتشفير النصوص الإلكترونية بطريقة موحدة أو معيارية، لتبادل المدونات بين الباحثين ولأغراض مختلفة، مثل استرجاع المعلومات والنشر الإلكتروني، مكتوبة بلغة الإس جي إم إل SGML.

تي إي آي: TEI

اختصار لمصطلح مبادرة تشفير النصوص Text Encoding Initiative.

استخراج المصطلح: term extraction

تحليل النص أو المدونة بغرض استخراج المصطلحات منها.

بنك المعطيات المصطلحية: terminological data bank (TDB)

بنك للمصطلحات والمعلومات المتعلقة بها، مثل التعريفات والمقابلات في لغات مختلفة.

نظام معالجة المصطلحات: terminological management system

الأدوات المستعملة في تسجيل البيانات أو المعطيات المصطلحية وترتيبها واسترجاعها بطريقة علمية منهجية.

قواعد المعطيات المصطلحية: terminology databases

قواعد معلومات تشتمل على المصطلحات وكل مايتعلق بها من معلومات، غالبا بطريقة مفصلة، و تستخلص منها بنوك المصطلحات. هناك من يرى بأن بنك المصطلحات قد يشتمل على عدد من قواعد المعطيات المصطلحية.

تقسيم النص: text chunking

مصطلح يطلق على تقسيم الجمل إلى فقرات في ضوء تحليل مبدئي.

نوع النص: text type

التخصص الذي ينتمي إليه النص، مثلا: أدبي أو علمي...

كلمة فعلية token:

الكلمة الواردة في النص أو المدونة بأية صورة (تصريف، مثل: كتبوا، يكتبون، نكتب، كتبا...و قلم، قلمان، أقلام...).

بنك الشجرات التركيبية: treebank

مصطلح يطلق أحيانا على المدونات المعرَبَة. ويأتي المصطلح من كلمة tree وهي الشجرة في التحليل النحوي الذي يبدأ بالجملة. ثم مايتفرع منها من مكونات، ويمكن أن نمثلها بالشجرة المقلوبة

كلمة نوعية: type

الكلمة الأساس التي ترد في أشكال مختلفة، وتعتبر هذه الصور المختلفة أمثلة للكلمة النوعية نفسها. ويقابلها الكلمة الفعلية token.

نسبة الكلمة النوعية للكلمات الفعلية::type-token ratio

العلاقة التناسبية بين الكلمة الواحدة وعدد صورها التي وردت في النص. ويعتبر النص غنيا بالكلمات كلما قلت هذه النسبة، حيث إن ذلك يعد مؤشرا على استخدام المؤلف لعدد أكبر من الألفاظ، بدلا من تكرارنفس الألفاظ.

مدونة غيرمحشاة: unannotated corpus

يطلق هذا المصطلح على المدونة الخام raw كما وردت في الأصل،والتي لم يضف إلى نصوصها أية سمات.

المحرف غير المحدد::wildcard

المحرف character، مثل النجمة أوعلامة الاستفهام (* ؟)،الذي يستعمل بدلا من أي محرف أوأكثر عند البحث عن الأنماط المختلفة (مثل ب* بمعنى ب متبوعا بأي عدد من الحروف، ليورد لنا الحاسوب جميع الكلمات التي تبدأ بحرف الباء).

قائمة الكلمات:wordlist

الكلمات الواردة في النص أو المدونة، غالبا مع معلومات عن مرات تكرارها.

ثانيا: القائمة العربية-الإنجليزية

لمعرفة معاني المصطلحات، انظر المسرد المشروح أعلاه والمرتب حسب الألفاظ الإنجليزية.

إحصاء دلالي: semantic count

أدوات (بحث) مشاعة: public domain tools

أدوات استعلام: query(ing) tools

أدوات اللسانيات الحاسوبية: CL tools

إرشادات التي إي آي: TEI Guidelines

إزالة الغموض: disambiguation

إس جي إم إل: SGML Markup Language

استخراج المصطلح: term extraction

إعراب كامل:full parsing

إعراب جزئي: skeleton parsing

إعراب: parsing

إمكانية الاستعادة:recoverability

بحث بالسياق: context search

بنك المعطيات المصطلحية: terminological data bank (TDB)

بيانات عن (شيء ما) :metadata

تجريد الكلمة: lemmatization

تحشية الضمير العائد: anaphoric annotation

تحشية خطابية: discoursal annotation

تحشية دلالية: semantic annotation

تحشية/ تعليقات: annotation

التحليل التحاوري: interactive analysis

ترتيب: sort(ing)

ترويسة المدونة: corpus header

ترويسة: header

تسوية: normalization

تشفير المحارف: character encoding

تشفير: encoding

تعليم النص: mark-up

تقسيم النص: text chunking

تلازم نحوي: colligation

تمثيل (اللغة): representativeness

توازن المدونة: corpus balance

تي إي آي: TEI

جدولة مقطعية:cross-tabulation

.جمع compilation:

دي تي دي: DTD

ذكاء اصطناعي: Artificial Intelligence (AI)

زود النص بالتحشية: annotate

سلسلة::string

سمة: tag

سي إي إس: CES

سياق غني بالمعلومات: knowledge-rich context

سياق: context

شيوع/ تكرار frequency:

صاحب لفظيا: collocate with

lexicographyصناعة المعاجم/ المعجمية:

طقعم أساس للواسمات: base tagset

طقم السمات:tagset

عنصر هدف/ عقدة: node

قائمة الاستبعاد: stop list

قائمة ألفبائية: alphabetical list

قائمة الكلمات:wordlist

قاعدة بيانات معجمية: lexical data base (LDB)

قواعد المعطيات المصطلحية: terminology databases

كتابة صوتية:phonetic transcription

كشاف سياقي ثنائي اللغة: bilingual concordance

كشاف سياقي: concordance

كلمة أساس:lemma

الكلمة المفتاحية: .keyword

كلمة فعلية token:

بنك الشجرات التركيبية: treebank

كلمة نوعية: type

كويك (الكلمة المفتاحية في سياق: KWIC (Key-Word In Context)

اللسانيات الحاسوبية: Computational Linguistics (CL)

لسانيات المدونات (اللغوية): Corpus Linguisticاللغة لأغراض خاصة: LSP

اللغة للأغراض العامة: LGP

محاذاة: alignment

المحرف غير المحدد: wildcard

محرك بحث: search engine

مدقق أسلوبي:style checker

مدقق إملائي: spelling checker

مدقق عام للنص: general text checker

مدونات: جمع مدونة.:corpora

مدونة أحادية اللغة: monolingual corpus

مدونة المتعلمين: :learner corpus

مدونة بها تحشية: annotated corpus

مدونة تاريخية: diachronic corpus

مدونة تزامنية::synchronic corpus

مدونة خاصة: special corpus

مدونة خام: raw corpus

مدونة عينة:sampled corpus

مدونة غيرمحشاة: unannotated corpus

مدونة فرعية::subcorpus

مدونة متشابهة: comparable corpus

مدونة متعددة اللغات: multilingual corpus

مدونة متوازنة: balanced corpus

مدونة متوازية (محاذاة): parallel (aligned) corpus multilingual.

مدونة محكية: spoken corpus

مدونة مراقبة (التطور): monitor corpus

مدونة مشبّعة: saturated corpus

مدونة معرَبَة: parsed corpus

مدونة مقارنة: comparative corpus

مدونة:corpus

مدونةمرجعية: reference corpus

مدى: span

مرتِّب حسب اليسار: left sorted

مرتِّب حسب اليمين: right sorted

مصاحب لفظي: collocate

مصاحبة لفظية: collocation

مصطلح محتمل: candidate term

مطابَقة الواقع: authenticity

معالجة اللغات الطبيعية: Natural Language Processing (NLP)

معالجة المدونة: corpus processing

معجم: lexicon

معرِب (نحوي): parser/syntactic parser

lexical resources/dataمصادر/ بيانات معجمية:

semantic prosodyالمعنى الناتج:

موارد/ مصادر لغوية: language/linguistic resources

نسبة الكلمة النوعية للكلمات الفعلية::type-token ratio

نص: co-text

نظام معالجة المصطلحات: terminological management system

نمط بحث: search pattern

نوع النص: text type

هندسة لغوية: language engineering

واسم لأقسام الكلام: part-of-speech tagger

وحدة معجمية: lexeme

وسم الأخطاء::error tagging

وسم بنفس السمة: ditto tagging

وسم مطابق:ditto tag

وسم موجه نحو المشكلة: problem-oriented tagging

وسم: tagging

خاتمة الكتاب

اللغة العربية والتقنيات الحديثة

المقدمة:

من المعروف أن العلاقة بين اللغة العربية والتقنيات الحديثة ترجع إلى أكثر من نصف قرن من الزمان وهي بين مد وجزر، معظمه مد خاصة في القرن الحادي والعشرين. ونأمل أن تتوطد العلاقة وتستفيد اللغة العربية والعرب من التقنيات الحديثة في خدمة اللغة التي تحتل المرتبة الرابعة أو الخامسة في سلم لغات العالم نتيجة لعوامل عديدة لا مجال لذكرها.

ما المقصود بالتقنيات أو التكنولوجيا الحديثة؟ يبدو أن معظم الباحثين يتفقون أن هناك صور عديدة لهذه التقنيات. ويمكننا أن نلخصها فيما يلي:

- الحاسوب باستخداماته الفردية stand alone، مثل الحاسوب الشخصي وأشباهه من الأجهزة اللوحية tablets.
- البرمجيات الحاسوبية المختلفة الأساسية والتطبيقية، والتي تتدرج من برمجيات معالجة النصوص والتدقيق الإملائي والنحوي والتعرف على اللغة المنطوقة أي الكلام speech recognition وتوليف الكلام speech synthesis وانتهاء بالذكاء الاصطناعي artificial intelligence (AI) وتطبيقاته التي بدأت تغزو شتى مجالات الحياة المعاصرة.
- الشابكة (الإنترنت).
- الهواتف الذكية.

في ضوء ما سبق يتضح لنا مدى اتساع الموضوع الذي نحن بصدد الحديث عنه. ولكن لمحدودية الوقت المتاح لهذه الورقة، فسنعرج إلى أهم المجالات التي تخدم فيها التقنيات الحديثة اللغة العربية، وتشمل: بعض البرمجيات الأساسية والتطبيقية المعروفة، دون الخوض في تفاصيلها.

١- أهم المجالات التي تخدم فيها التقنيات الحديثة اللغة العربية:

١-١- البرمجيات الأساسية:

تشمل هذه البرمجيات التي تعتبر ضرورة لكثير من التطبيقات الحاسوبية في خدمة اللغة العربية والناطقين بها. ومن هذه

١-١-١- برمجيات التعرف على الكلام. فهذه مطلوبة في تطبيقات الاستعلام الشفوي والتعليمات المنطوقة والترجمة الشفوية.

١-١-٢- برمجيات توليف الكلام أي النطق، وهي ضرورة في تطبيقات مثل القراءة الشفوية للنصوص (لمساعدة المكفوفين وضعاف البصر مثلًا) وتقديم الإرشادات الشفوية (كما في التعليمات التي تقدمها خرائط قوقل google maps وما شابهها، وكذلك في الترجمة الآلية الشفوية.

١-١-٣- برمجيات التعرف على النصوص المكتوبة يدويًا أو المطبوعة (مثل التعرف على المحارف ضوئيًا cognition (OCR) optical character re وتحويلها إلى نصوص رقمية. وهذه مطلوبة في مجالات تطبيقية كثيرة مثل حفظ الوثائق وأرشفتها حاسوبيًا، وفي إعداد المدونات اللغوية corpora التي سنعرج عليها لاحقًا. هناك برمجيات أعدت لهذا الغرض في الساحة العربية، منها مثلًا "القارئ العربي" لتحويل النصوص المطبوعة إلى نصوص رقمية قابلة للتعديل عليها.

١-١-٤- برمجيات تحليل النصوص صرفيًا ونحويًا (الإعراب الآلي). وهذه مطلوبة في مجالات كثيرة من التطبيقات اللغوية مثل فهم النصوص والترجمة الآلية والتشكيل الآلي للنصوص العربية، والدراسات المعجمية الإحصائية للنصوص وكذلك الإحصاء التركيبي لها.

ويسعدنا أن نقول أن هناك عددًا من هذه البرمجيات المتاحة في الساحة العربية، وقد بني عليها برمجيات (للتشكيل الآلي) مثل: برامج (حركات) و(تشكيل) و (مشكال).

١-٢- تطبيقات التقنيات الحديثة في خدمة اللغة العربية:

لا شك أن تطبيقات التقنيات الحديثة كثيرة لا يمكن حصرها في دقائق ولا ساعات ربما. لذلك سأكتفي بذكر بعض منها:

١-٢-١- صناعة المعاجم العربية:

تمثل صناعة المعاجم مثالًا جيدًا لأوجه الاستفادة من الحاسوب في خدمة اللغة العربية. ويشمل ذلك مراحل العمل المختلفة في إعداد المعاجم العربية:

- اختيار مداخل المعجم وفق أسس علمية منهجية، مثل الاستفادة من المدونات اللغوية الكبيرة في توفير النصوص التي نختار منها مداخل المعجم بصورة واقعية. كذلك تفيدنا بعض البرمجيات المرتبطة بالمدونات مثل المكشاف السياقي concrdancer في إجراء تحليل إحصائي للألفاظ والتعابير الواردة في المدونة وترتيبها بحسب تكرارها، لنختار الأشيع منها مثلًا في إعداد المعجم، كما سنبين لاحقًا عند الحديث عن المدونات اللغوية. وهناك عدد من المعاجم العربية التي اعتمدت المدونات اللغوية المحوسبة في اختيار مداخلها، مثل المعاجم التي أصدرتها شركة صخر وأعدها مختار عمر (بالتعاون مع فريق من الباحثين) ووقاء كامل فايد، مما نذكر بالتفصيل لاحقًا.
- إعداد المعجم، من حيث مداخل ومكونات كل مدخل (معلومات صرفية كالجذر، نحوية، إملائية، تعريفات، أمثلة للاستعمال، مصاحبات لفظية، تعابير اصطلاحية، مرادفات، أضداد...)، حيث يخصص سجل record لكل كلمة أو مدخل، ويخصص لكل من هذه المكونات حقل field في برنامج إدارة البيانات (قواعد البيانات).
- ويتم تحويل هذه البيانات إلى نصوص طبيعية عند نشر المعجم في صورة ورقية مطبوعة أو رقمية.

- تقديم المادة المعجمية:
- من مميزات استخدام الحاسوب في تقديم المادة المعجمية، إمكانية تقديمها بصورة تفاعلية، تعتمد على حاجات الباحث ورغباته.
- يتيح لنا الحاسوب البحث عن الكلمة في صورها المختلفة مع إيراد المعلومات المتعلقة بها.

ولعل من أمثلة الاستفادة من الحاسوب في صناعة المعجم العربي "المعجم العربي التفاعلي" في نسخته الثالثة. وهو منتج عربي شارك فيه عدد من المنظمات العربية (المنظمة العربية للتربية والثقافة والعلوم ومدينة الملك عبدالعزيز للعلوم والتقنية ومؤسسات علمية أخرى في المغرب وسوريا.)

ومن أشهر الأمثلة لاستخدام الحاسوب في العمل المعجمي معجم "المعاني" المتاح بصور مختلفة على الشابكة، وكذلك بنوك المصطلحات الألية.

ومن أمثلة بنوك المصطلحات العربية، مثل "باسم (البنك الآلي السعودي للمصطلحات) وبنك "معربي" التابع لمعهد الدراسات والبحوث للتعريب التابع لجامعة محمد الخامس، وبنك المصطلحات التابع لمكتب تنسيق التعريب في الوطن العربي في الرباط، وبنك أراب تيرم Arabterm الذي أنشأه المكتب المذكور بالتعاون مع جهات أخرى مولتها الحكومة الألمانية. فهذه كلها تتيح للباحث البحث عن المصطلح بأي من اللغات المتاحة في بنك المصطلحات (العربية، الإنجليزية، الفرنسية، الألمانية).

- يجدر بالذكر أن هناك عددًا من المعاجم العربية التراثية (مثل لسان العرب) والحديثة العامة والمتخصصة (مثل المعجم الطبي الموحد) المتاحة على الشابكة.

١-٢-٢- المدونات اللغوية العربية:

تمثل المدونات اللغوية العربية Arabic corporaسواء ما أنجزته هيئات عربية مثل مدونة اللغة العربية من إعداد مدينة الملك عبدالعزيز للعلوم والتقنية في الرياض، والمدونة العربية الدولية الذي أنجزته مكتبة الاسكندرية، أم أنجزه مستشرقون، مثل Arabicorpus ومدونة القرآن الكريم Quran-corpus ذخائر قيمة جدًا لأغراض البحث اللغوي العربية، مثل صناعة المعاجم وتطوير البرامج اللغوية الحاسوبية computational linguistics، مثل البرمجيات التي وصفناها بالبرمجيات الأساسية، كما أن المدونات ثنائية اللغة المتوازية parallel corpora تعتبر منطلقًا لتطوير بعض أنظمة الترجمة الآلية المبنية على النظام الإحصائي، إلى جانب إعداد المعاجم ثنائية اللغة. ولا نشك في فائدة المدونات اللغوية في تطوير برمجيات الذكاء الاصطناعي في مجالات عديدة، مثل فهم اللغة وإنتاجها والترجمة الآلية، إضافة إلى فائدتها في عدد من المجالات مثل اللسانيات الاجتماعية وتعليم اللغة لأهلها ولغير الناطقين باللغة العربية.

هذا وتتوافر في الساحة العشرات من هذه المدونات بأنواعها العامة والخاصة (مثل القرآن الكريم)، كما تشتمل على مدونات للغة العربية الفصيحة والفصحى وبعض العاميات كذلك. وتختلف هذه المدونات اللغوية من حيث محتوياتها وأحجامها وكمية المعلومات المتوافرة بها.

وقد أحصينا حتى تاريخه أكثر من خمسين مدونة لغوية للغة العربية ما بين عامة وخاصة وبين فصيحة وعامية، أسهم في إعدادها مؤسسات عربية (مثل مدينة الملك عبدالعزيز للعلوم والتقنية في الرياض ومكتبة الإسكندرية) ومؤسسات أجنبية (مثل جامعة بريغهام يونغ في الولايات المتحدة ولوفان في بلجيكا وعدد من ا لمدونات الخاصة بالقرآن الكريم والعامة التي أعدها باحثون في جامعة ليدز في بريطانيا). ويتراوح عدد الكلمات في هذه من المليارات إلى مئات الآلاف من الكلمات.

(لمزيد من المعلومات عن المدونات اللغوية العربية وكيفية الاستفادة منها في المجالات المختلفة، انظر: محمود إسماعيل صالح "المدونات اللغوية العربية وطرق الإفادة منها")

- وقد استفاد بعض الباحثين من المدونات المختلفة في إعداد معاجم عربية حديثة مبنية على لغة الاستعمال الحقيقي، مثل (معجم العربية المعاصرة) الذي أعده أحمد مختارعمر بمعاونة فريق من الباحثين، و(معجم

التعابير الاصطلاحية) الذي أعدته وفاء كامل، و(المعجم العربي للطالب - لمراحل عمرية مختلفة) الذي أعده مجموعة من الباحثين السعوديين بتمويل من مدينة الملك عبدالعزيز للعلوم والتقنية ووزارة التعليم السعودية). ومنها (**المعجم المعاصر**) من إنتاج شركة صخر الكويتية. و**المعجم العربي للطلاب** (للمراحل الدراسية المختلفة) الذي أعده مجموعة من الباحثين السعوديين، كما يجدر بالذكر أحد المعاجم الرائدة المبنية على المفردات الشائعة، وهو **معجم الطالب السياقي** الذي بني على دراسة لداود عبده لأشيع ثلاثة آلاف كلمة في اللغة العربية.

وجدير بالذكر المعجم العربي الإنجليزي الذي أعده المستشرقان الأمريكيان بكوولتر وباركنسون، حيث أوردا الخمسة آلاف كلمة الأكثر شيوعًا في مدونة للغة العربية تجاوز عدد كلماتها ثلاثين مليون كلمة، مرتبة حسب تكرارها ومزودة بمقابلاتها الإنجليزية، إضافة إلى المعلومات المعجمية الأخرى.

أخيرًا لابد لنا أن ننوه بمشروع المعجم التاريخي للغة العربية الذي يعمل عليه اتحاد المجامع العربية، وقد صدر منه عدد من المجلدات لأحرف المعجم الأولى في العام ٢٠٢٠م، وكذلك معجم الدوحة التاريخي للغة العربية.

١-٢-٣- **تعليم اللغة العربية**:

تزخر المكتبة الإلكترونية بأنواع من البرامج والمنصات والمراجع المتاحة على الشابكة أو القابلة للاستخدام الشخصي في مجالات تعليم اللغة العربية للناطقين بها وبغيرها، خاصة الأخيرة التي أحصينا منها أكثر من ثمانين موقعًا تعلم اللغة العربية للناطقين بغيرها من إنتاج مؤسسات تعليمية أو شركات خاصة، وتتفاوت في شمولها وجودتها. هذا بالإضافة إلى أن هناك عددًا من المنصات التي أنشئت حديثًا في ظل جائحة كورونا وكذلك الكتب والسلاسل المتاحة في صورة رقمية ومتاحة على الشابكة (الإنترنت).

١-٢-٤- **الترجمة الألية**:

تعتبر الترجمة الآلية بشقيها المكتوب والشفوي من أشهر التطبيقات المعروفة للحاسوب والتقنيات المرتبطة به كالهواتف الذكية. وهناك عدد من برمجيات الترجمة بين العربية وعدد من اللغات، لعل من أشهرها google translate, Babylon ويبدو أن الأخيرة هي التي يوفرها برنامج Microsoft Word لمستخدميه. هناك

أيضًا برامج توفرها شركات عربية مثل صخر والناقل العربي Cimos والمترجم العربي ومشتقاتها على الشابكة مثل عجيب والمسبار، إضافة إلى عدد من الأنظمة التي تعمل على الهواتف الذكية.

١-٢-٥- **معينات الترجمة الآلية**:

لعل من أهم هذه المعينات أنظمة "ذاكرة الترجمة" translation memory وهو برنامج يتيح لمستخدمه تخزين النصوص المترجمة بصورة تتيح له الاستفادة منها تلقائيًا عند ترجمة نصوص جديدة مشابهة، إضافة إلى إعداد مسارد ثنائية اللغة. وتفيد ذاكرات الترجمة في تطبيقات الذكاء الاصطناعي في الترجمة، حيث تمثل نماذج يمكن للحاسوب التعلم منها.

- المعاجم الإلكترونية بصورها المختلفة: أحادية اللغة وثنائيتها ومعاجم المترادفات.

(لمزيد من المعلومات حول هذا الموضوع، انظر محمود إسماعيل صالح "التقنية في خدمة الترجمة والمترجمين".)

المراجع

رشوان، محسن والسعيد، المعتزبالله (٢٠١٩) **مقدمة في حوسبة اللغة العربية**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

صالح، محمود إسماعيل (٢٠٢٠) "الحاسوب في تعليم اللغات" في مجلة: **تعليم العربية لغة ثانية**، السنة الثانية، العدد الرابع، صص ٧-٤٤.

صالح، محمود إسماعيل (٢٠٢٠) **دليل السلاسل التعليمية والمواقع الإلكترونية لتعليم العربية للناطقين بغيرها**. الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

صالح، محمود إسماعيل (٢٠١٨) "التقنية في خدمة الترجمة والمترجمين" في عمرالمهديوي (تحرير): **اللسانيات الحاسوبية واللغة العربية: إشكالات وحلول.** عمان: دار كنوز المعرفة، صص ٥٣-٨٦.

صالح، محمود إسماعيل "الترجمة الآلية واللغة العربية". محاضرة متاحة على اليوتيوب.

صالح، محمود إسماعيل (٢٠١٥) المدونات اللغوية العربية: بناؤها وطرق الإفادة منها" في العصيمي، صالح (تحرير، ٢٠١٥) **المدونات اللغوية: بناؤها وطرق الإفادة منها.** الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية، صص ١٨-٧٧.

صالح، محمود إسماعيل."لسانيات المدونات اللغوية: مقدمة للقارئ العربي" في:مدونة الدكتور محمود إسماعيل صالح على الشابكة.

صالح، محمود إسماعيل (٢٠١٩) "بنوك المصطلحات الآلية" ورقة مقدمة إلى **ورقة عمل مقدمة إلى مؤتمر الدورة الخامسة والثمانين لمجمع اللغة العربية بالقاهرة.**

العصيمي، صالح (تحرير، ٢٠١٥) **المدونات اللغوية: بناؤها وطرق الإفادة منها.** الرياض: مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية.

المهديوي، عمر (تحرير، ٢٠١٨): **اللسانيات الحاسوبية واللغة العربية: إشكالات وحلول.** عمان: دار كنوز المعرفة.

محتويات الكتاب

المقدمة ٣

الفصل الأول ٥

الحاسوب واللغة ٥

١- الاستعمال العام للحاسوب في النشاط اللغوي: ٥

١- ١ - معالجة النصوص أو تنسيق الكلمات word processing: ٥

١-٢- عرض النصوص لأغراض الشرح والمناقشة الجماعية والتدريب (برنامج power point مثلا)..... ....٥

١-٣-استخدام الشابكة (الإنترنت) لأغراض مختلفة ٥

١-٤-التعليم والتقويم: ٦

١-٥-الترجمة: ٦

١-٦- التعرف على النصوص المكتوبة: ٦

١-٧-توليف الكلام: ٦

١-٨-التعرف على النصوص الشفوية: ٧

٢- البحث الغوي: ٧

٢-١- توليف الكلام والبحث اللغوي: ٧

٢-٢- لسانيات المدونات اللغوية corpus linguistics: ٧

٢-٣- قواعد البيانات والبحث اللغوي: ٧

٣- اللسانيات الحاسوبية: ٨

٤- مجالات بحثية غير واضحة المعالم: ٩

الفصل الثاني التعليم و التقويم.....١٨

١-التعليم والحاسوب: ١٩

٢-تعليم اللغة: ٢٠

٢-١-تعليم عناصر اللغة: ٢٠

٢-٢ تعليم مهارات اللغة ٢٨

٢-٣ مثال على حزمة برامج شاملة لتعليم اللغة العربية: ٣٨

٢-٤التأليف وتصميم الاختبارات: ٣٨

٢-٥- استخدامات الحاسوب في الاختبارات: ٣٩

٢-٦- تقويم المواد التعليمية: ٤٢

٢-٧- الشابكة (الإنترنت) في تعليم اللغة: ٤٢

٢-٨- بعض مشكلات تعليم اللغة باستخدام الحاسوب وحلول مقترحة ٤٤

الفصل الثالث الترجمة والتعريب..... ..... ٥٦

التقنية في خدمة الترجمة: طرق الاستفادة والتقويم ٥٧

١,١. الترجمة الآلية: ٥٧

٢,١. معيبنات حاسوبية للترجمة: ٦٤

٣,١. المعينات المعجمية في الترجمة: ٦٦

١-٤- معينات حاسوبية أخرى: ٦٩

٢- تقويم أنظمة الترجمة الآلية: ٧١

٢-١- تقويم الترجمة: ٧٢

٢ -٢- نظام BLEU الآلي لتقييم الترجمة الآلية: ٧٩

٢-٣- مقترحات لتنقيح النصوص المترجمة آليا: ٧٩

الفصل الرابع: ٨٥

الذكاء الاصطناعي والترجمة من العربية وإليها ٨٥

١- مشكلات الترجمة الآلية من العربية المكتوبة وإليها مع نماذج حقيقية:..... ٨٥

١-١-مشكلات الترجمة الآلية من العربية المكتوبة: ٨٥

١-٢-التذكير والتأنيث للكلمات التي لاتفرق بينها بعض اللغات (مثل teacher .....).....٨٨

١-٣- قواعد المطابقة: ٨٨

١-٤- هذا إلى جانب المشكلات العامة بين اللغات: مثل تعدد المعاني، الاشتراك اللفظي، التعبيرات الاصطلاحية، الاستعمال المجازي للألفاظ والعبارات، الإحالة reference) ٨٨

٢- أساليب التعامل مع المشكلات المشار إليها أعلاه:..... ٨٩

٢-١- التحرير المسبق للنصوص..... ٨٩

٢-٢- الترجمة التحاورية، حيث يتم التعاون بين المترجم والآلة. ٨٩

٢-٣- التحرير اللاحق. ٨٩

٢-٤- الأسلوب الإحصائي ٨٩

٢-٥- الذكاء الاصطناعي: ٩٠

٣- دور الذكاء الاصطناعي..... ....٩٠

٣-١- أنظمة الترجمة الحديثة.....٩٠

الفصل الخامس ٩٧

التعريب الآلي للأسماء الأجنبية: الأسس اللغوية ٩٧
١- القضايا والأسس العلمية اللغوية لتعريب الأسماء: ٩٧
١-١- الكتابة الصوتية أم النقحرة: ٩٧
١-٢- ما هي أسس النقل؟ ٩٨
٢- مشكلات التعريب للأسماء المكتوبة بالحرف اللاتيني: ٩٩
٣- أهم مشكلات كتابة الأسماء الواردة في لغات أوربا الغربية: ١٠٠
٤- مشكلات تعريب الأسماء الأجنبية: ١٠٢
٥- الحلول المقترحة والمتبعة للمشكلات المذكورة: ١٠٣
٥-١- الأسلوب الإحصائي: ١٠٣
٥-٢- الأسلوب اللغوي الجغرافي: ١٠٤
٦- أساليب المراجعة: ١٠٥
٦-١- أهم أسس التقويم والمراجعة: ١٠٦
الفصل السادس صناعة المعاجم العامة والمتخصصة (المصطلحات) ١٢٤
الحاسوب في العمل المعجمي ١٢٤
١- أنواع المعاجم: ١٢٤
١-١- الجذر..... .... ١٢٦
١-٢- نظام التقفية..... ..... ١٢٦
١-٣- نطق الكلمة في أبسط صورها بعد تجريدها من الإضافات التصريفية: ١٢٦
١-٤- الموضوعات..... ..... ١٢٦

٢- الخطوات المتبعة في إعداد المعاجم: ١٢٧
٢-١- تحديد الهدف والجمهور المستهدف: ١٢٧
٢-٢- تحديد مداخل المعجم ومحتويات المدخل: ١٢٧
٢-٣- إعداد المعجم ونشره لتقديمه لجمهور المستفيدين. ١٢٨
٣- أهم استخدامات الحاسوب في العمل المعجمي. ١٢٨
٣-١- المدونات اللغوية: ١٢٨
٣-٢- تحديد الألفاظ الشائعة: ١٢٩
٣-٣- تحديد مفهوم الكلمة والمصطلحات ذات العلاقة: ١٢٩
٣-٤- الأدوات اللازمة للاستفادة من المدونات اللغوية: ١٣١
٣-٥- مجالات البحث التي تتعلق بالجانب المعجمي للغة: ١٣١
٤- صناعة المعاجم العربية: ١٣٣
٤-١ اختيار مداخل المعجم ومكوناتها: ١٣٣
٤-٢ - إعداد المعجم: ١٣٣
٤-٣- تقديم المادة المعجمية: ١٣٤
الفصل السابع..... .....١٤١
التقنية في خدمة المصطلح العلمي العربي..... ١٤١
١ -المدونات اللغوية: ١٤١
٢ - قواعد بيانات المصطلحات التراثية: ١٤١
٣ - قواعد بيانات المصطلحات العلمية الحديثة: ١٤٢

٤ -لبنوك الآلية للمصطلحات العلمية: ١٤٢
٤-١ -بنوك المصطلحات العلمية المحلية: ١٤٣
٤--٢بنك المصطلحات العلمية المجمعي: ١٤٣
٤-٣- البنك العربي المركزي للمصطلحات العلمية: ١٤٣
٥- منصة المصطلح العلمي العربي: ١٤٤
٦- التدريب الافتراضي: ١٤٥
الفصل الثامن بنوك المصطلحات الآلية..... ١٤٧
١- مميزات بنوك المصطلحات الآلية: ١٤٨
١-١- أهم الفروق بين بنك المصطحات والمعجم التقليدي: ١٤٨
١-٢- أهم ميزات بنك المصطلحات: ١٤٨
١-٣- المستفيدون واحتياجاتهم: ١٥٠
الفصل التاسع البحث اللغوي (مثال لسانيات المدونات اللغوية) ١٦٦
المدونات اللغوية ولسانيات المدونات..... ١٦٦
١- تعريف المدونة اللغوية ولسانيات المدونات: ١٦٦
٢- مقارنة بين لسانيات المدونات والتحليل اللغوي التقليدي: ١٦٧
٣- مزايا لسانيات المدونات (المحوسبة): ١٦٨
٤- أنواع المدونات اللغوية: ١٦٩
٥- مواصفات المدونات اللغوية: ١٧٢
٦- إعداد المدونة و طرق الاستفادة منها: ١٧٩

٧- مصطلحات مهمة في مجال لسانيات المدونات: ١٨٢
٨- أساليب الوسم المختلفة: ١٨٤
٩- أمثلة لمدونات للغة العربية..... ١٨٥
١٠- متطلبات التعامل مع المدونات: ١٩١
١١- مجالات الإفادة من المدونات اللغوية: ١٩٣
الخاتمة ٢٦٢
المراجع..... ٢٦٩
محتويات الكتاب.....٢٧٠