مؤتمر التعريب الحادي عشر
المنظمة العربية للتربية والثقافة والعلوم
عمان ١٢-١٦/١٠/٢٠٠٨

اللغة العربية والترجمة الآلية
المشاكل والحلول

أ.د. محمد زكي خضر
الجامعة الأردنية

الخلاصة

يهدف هذا البحث إلى إعطاء فكرة عن تطور الترجمة الآلية بين اللغات العالمية وآخر ما وصل إليه التقدم في هذا المجال وخاصة في الترجمة الأحصائية. يناقش البحث الوضع بالنسبة للترجمة الآلية من اللغة العربية وإليها والمشاكل التي تعاني منها ومن ثم تبيان الخطوات التي ينبغي اتباعها للتقدم في هذا المجال. تكون المعاجم والمصطلحات ركيزةً رئيسيةً في الترجمة كما تكون الذخيرة اللغوية حديثًا والعمليات الإحصائية عليها جزءًا مهمًا من مرتكزات الترجمة الآلية. لذلك فإن إعداد معجم عربي محوسب وذخيرة لغوية متعددة اللغات تعتبر خطوات هامة في طريق تطوير الترجمة الآلية من اللغة العربية وإليها. يتطرق البحث أيضًا إلى مختلف جوانب الترجمة الآلية كالترجمة الفورية والترجمة متعددة اللغات والبرامج المساعدة للترجمة البشرية والترجمة بمساعدة الحاسوب وإلى وضع الترجمة على شبكة الإنترنيت.

يخلص البحث إلى ايضاح المشاكل التي تعاني منها الترجمة الآلية من اللغة العربية وإليها اليوم ويشير إلى الحلول التي تؤدي في المستقبل إلى الإفادة من الترجمة الآلية بنطاق واسع.

مقدمة

الترجمة هي نقل معاني نص من لغة إلى لغة أخرى مع مراعاة الدقة والأسلوب . ويتطلب ذلك فهم النص الأصلي والتعبير عن المحتوى والأسلوب بلغة أخرى. فالمترجم يجب أن يتقن اللغتين االمترجم منها والمترجم إليها.

هناك طريقتان معروفتان في الترجمة. أولاهما تعتمد الترجمة الحرفية والإلتزام بمعاني مفردات النص الأصلي ونقلها إلى اللغة الثانية والطريقة الثانية تعتمد على فهم المعنى العام ثم التعبير عنه باللغة الثانية بأسلوب المترجم نفسه.

المترجم يتعرف على الرموز المكتوبة في الترجمة الكتابية والأصوات المنطوقة في الترجمة الشفوية ، أي يقرأ الرموز الكتابية للغة التي يترجم منها إذا كان النص مكتوبًا أو يتعرف على أصوات اللغة التي يستمع اليها إن كانت ترجمة شفوية. وهنا نذكر أن قراءة اللغة العربية المكتوبة بحروف غير مشكولة يستوجب معرفة من القارئ لكي يقوم هو بالتشكيل الصحيح دون أن تبصر عيناه علامات التشكيل. وهو بذلك يستعمل خبراته التي تعلمها وتدرب عليها في ماضي حياته من قواعد اللغة لكي يستعملها بشكل سريع فينصب المفعول ويرفع الفاعل

وربما يفعل ذلك بالسليقة وهو لا يعرف من قواعد إعراب اللغة العربية شيئًا. أما المترجم للكلام المنطوق فهو يستطيع تمييز الجملة التي يترجمها إن كانت إستفهامية أم خبرية أم تعجبية دون وجود علامات استفهام أو تعجب. بعد ذلك يرجع المترجم إلى الوحدات المعجمية وهي الكلمات والتعبيرات الاصطلاحية ويفهم معانيها في سياقاتها اللغوية والاجتماعية المختلفة ، وهنا يجب التأكيد على السياقات المختلفة، لأن الكلمة الواحدة تعني أشياء كثيرة طبقا للسياق الذي ترد فيه. هناك إذن مشكلة معاني الكلمات في السياقات المختلفة (أي تعدد المعاني) ، إضافة إلى المشتركات اللفظية (الكلمات التي تتشابه في كتابتها أو نطقها وتختلف في معانيها، مثل عين الإنسان وعين الماء) (١) وهناك التعبيرات الاصطلاحية .وهذه كما سنرى تمثل مشكلة كبيرة بالنسبة للترجمة عامة والترجمة الآلية خاصة.

وهنا ينبغي تحليل المفهوم والوقوف على كنهه وما يتفرع عنه من ظلال في المعنى، وخاصة المجازية منها وغيرها، وذلك حتى تتحدد العلاقة بين هذا المفهوم الأساس وفروعه وبين المفاهيم التي يمكن أن تتصل به بشكل أو بآخر، أي العلاقة بينه وبين مجموعات المفاهيم الأخرى، وذلك للوصول في النهاية إلى الجملة أو شبه الجملة. فالمفهوم الذي يرتبط بجسم واحد بعينه، أو يدل عليه، هو مفهوم إفرادي، مثل "المريخ." أما إذا ارتبط المفهوم بعدة أجسام بينها نوع من التجانس الذي يضعها في مجموعة واحدة، فإن المفهوم يكون عاماً (بمعناها الضيق هنا)، مثل "كوكب." فهذا المفهوم يدل على عدد من الأجسام التي تدور حول الشمس. ولتحديد المقصود يمكن أن يكون التعريف مركِّزًا أو موسعًا . وعلى ذلك فإن دقة التعريف تحدد المفهوم. والتعريف قد يكون شاملاً في بعض الأحيان وفي غالب الأحيان لا يخلو من النقص أو أن يكون تعريفًا بالضد أو ما يعرف بالتعريف السلبي أو أن يكون فصفاضًا أو ضيقًا أو يأخذ بمبدأ الإحلال والتعويض إلى غير ذلك من الصفات التي يتصف بها التعريف. وهذا التعريف قد يكون واضحًا في ذهن المترجم وقد يكون مشوشًا فيتصرف في ترجمته لهذا المفهوم وقد يعبر عنه بمفهوم يفهمه هو ويكون قريبًا أو بعيدًا من المفهوم الأصلي

ومن العمليات التي تساعد في التعرف على المفهوم التعرف على الوحدات النحوية وعلى وظائفها : أي تحديد الوحدات (الألفاظ والتعبيرات) ومعرفة دور كل منها في الجملة ، مثل المسند والمسند إليه والتكملة . ومن أمثلة التعبيرات المضاف والمضاف إليه ، والجار والمجرور والصفة والموصوف وغير ذلك . وكل من هذه التعبيرات له وظيفته النحوية , فبعضها يقوم مقام الفاعل أو المفعول به وبعضها يقوم مقام الصفة أو الظرفية إلى غير ذلك . ومن الوحدات النحوية الجمل . ولابد للمترجم أن يفهم معنى كل واحد من هذه العناصر ووظيفته اللغوية والنحوية في النص . فصيغة الأمر قد تعبر عن إلزام السامع أو رجاءه أواستعطافه مثلا . كما أن الصيغة الخبرية قد تأتي للإخبار كما تستعمل للطلب غير المباشر(كما في عبارة "أرى أن لديك قلما زائدا") ، بل وللدعاء كذلك ("رحم الله فلاناً" مثلا). (٢)

ومن الأمور المهمة في الترجمة معرفة حقل النص ومجال تخصصه . مثلا إذا النص في علم الكيمياء فإنك تجد أن بعض الكلمات تختلف معانيها في هذا النص مما لو وردت في نص في الفيزياء أو وردت في الالكترونيات. أو أن كلمة وردت في حقل تجارة المواد الغذائية أو في حقل من حقول العلوم الزراعية. وهكذا.

الجانب الثاني وهوالتعبيرعن محتوى النص باللغة الهدف أي اللغة المترجم إليها. ويتطلب هذا معرفة بإنتاج المقابلات الصوتية أو الكتابية والنحوية والمعجمية والبلاغية . فعلى المترجم أن يعرف أن هذه التعبيرة الإسنادية مثلا تترجم إلى كذا وهو بذلك يجب أن يكون ملمًا بقواعد النطق أو الاملاء وأصول الكتابة : قواعد النطق إن كانت الترجمة شفوية في اللغة المترجم إليها ، أو قوعد الاملاء وأصول الكتابة في هذه اللغة، إن كانت الترجمة

تحريرية. وهو بذلك يقوم باختيار العبارات المناسبة في اللغة التي يترجم إليها التي توافق المفهوم الذي توصل إليه في الخطوات السابقة.

وهنا تبرز براعة المترجم في اللغتين حيث يجب أن يتقن ما يسمى بالتعادل أو التقابل المعجمي أو الاصطلاحي ، فالتعبيرة العربية "رجع بخفي حنين" ليس لها مقابل حرفي بأية لغة ، ولو ترجمت ترجمة حرفية لأصبحت مضحكة . فهو يجب أن يفهم مثل هذه العبارات ويجب أن يعرف العبارات الدارجة التي تقابلها في اللغة الثانية أو على الأقل يكوّن هو تعبيرًا من كلمات تلك اللغة بما يعبر عن المفهوم بدقة كافية مستعملاً القواعد النحوية والصرفية في اللغة التي يترجم إليها. فالترجمة كالكتابة أو التأليف تخضع لقواعد الصحة اللغوية نفسها.

أما القواعد الأسلوبية وهي القاعدة البلاغية المعروفة (لكل مقام مقال) فنجد أن الجملة أو العبارة قد تكون صحيحة نحويًا ومعجميًا ولكنها غير مناسبة من حيث المقام. فالمترجم يجب أن يكون ذا معرفة بالقواعد الأسلوبية للغتين المترجم منها والمترجم إليها. (٢)

وماذا بعد ذلك ؟ لا يزال هناك مشكلة التعامل مع الجمل الطويلة. فالجمل الطويلة هي من أهم مشاكل الترجمة لأنها تحتاج إلى معالجات ذهنية معقدة وقد تشوش ذهن المتلقي بشكل عام. فقد تحوي الجملة الطويلة على عدة مفاهيم وقد يتجزأ المفهوم الواحد ليتكامل بين بدء الجملة ونهايتها. وقد يتخلل المفهوم الواحد مفاهيم ثانوية. وقد يكون المفهوم المهم مخفيًا بين عدد من المفاهيم الأقل أهمية منه. وهكذا فإن الجملة الطويلة قد تستعصي على الفهم والتحليل وكذلك تستعصي على التركيب في اللغة المترجم إليها.

إن معلوماتنا عن الترجمة محدودة إلى الآن ، ولانستطيع أن نقول أن المترجم يتبع هذه الخطوات : ٣،٢،١ بل إن كل مترجم له أسلوبه الذي يتبعه في الترجمة ، والأصعب من ذلك تحديد كيفية فهم النص . ماذا نفعل تمامًا حينما نحاول فهم نص ما من النصوص ؟ إن الطريقة التي يتعامل بها دماغ الإنسان مع النصوص لا زالت غير معروفة على وجه الدقة وتخضع لنظريات غير مؤكدة لحد الآن . كما أنه لم تتوفر حتى يومنا هذا أساليب ناجحة للتعامل مع توفير عناصر الحس والخبرة الذاتية والمعلومات وغيرها التي يمكن أن ننقلها بشكل من الأشكال إلى الحاسوب في الترجمة الآلية التي سنتناولها في ما بعد . لذلك يمثل "فهم النص" مشكلة كبرى في الترجمة.

لماذا الترجمة الآلية؟

السبب الأول والأهم هو وجود كمية هائلة مما يجب ترجمته مما لا يكفي المترجمين من البشر القيام بجزء يسير منه خاصة بعد تفجر ثورة المعلومات وتنوع اللغات التي تنتج المعارف اليوم مما ينبغي معرفته ممن لا يتكلمون تلك اللغة. يزيد عدد اللغات الحية اليوم على ٤٠٠٠ لغة في كافة أنحاء العالم. وتأتي اللغة العربية ضمن اللغات العشرة الأولى منها إذا ما أخذنا بعين الإعتبار عدد الناطقين بها. ورغم التأريخ الحافل للغة العربية إلا أنها اليوم ليست المصدر الأهم في العلوم والتقنية الحديثة. لذلك فإن عملية الترجمة من اللغات الإخرى إلى اللغة العربية ذات أهمية بالغة بالنسبة للناطقين بالعربية. فإذا ما أرادوا الإطلاع على آخر ما توصل إليه العلم فإما ان يتعلموا لغة (أو لغات) أخرى بجانب العربية ، أو أن يترجم لهم ما يصدر من علوم في اللغات الأخرى. وما نشهده اليوم من إقبال على تعلم اللغة الإنكليزية والتدريس بها في الجامعات والمدارس في العالم العربي ما هو إلاّ نتيجة للتخلف في الترجمة إلى اللغة العربية.

أبرز تقرير التنمية الإنسانية العربية الأول لعام ٢٠٠٢ ضعف حركة الترجمة في الوطن العربي ، بل ضمورها الشديد ، مما يكاد يقضي على دورها المأمول في نقل المعرفة وتوطينها باللغة العربية. وقد كشفت الإحصاءات

في هذا الصدد أن ما يترجمه العالم العربي من كتب قد لا يزيد على خمس ما يترجمه بلد أوربي صغير مثل اليونان الذي يقل سكانه عن ٥% من سكان الوطن العربي. وتتعاظم أهمية الترجمة العلمية يومًا بعد يوم نتيجة الإنفجار المعرفي وتزداد هذه الأهمية بالنسبة لعالمنا العربي لكونه – أساسًا – متلقيًا للمعرفة أكثر منه منتجًا لها. وتعثر الترجمة يعد سببًا رئيسيًا وراء تعثر تعريب التعليم الجامعي . (٣)

لمقدم هذا البحث تجربة في ترجمة أحد الكتب العلمية التي تدرس في أقسام الهندسة الكهربائية في كافة أنحاء العالم مثذ السببعينيات في القرن الماضي. فبعد قضاء أكثر من عام في ترجمة الطبعة الثانية من الكتاب ظهرت الطبعة الثالثة فأعيدت الترجمة من البداية وأدى ذلك إلى تأخر ظهور الترجمة العربية أكثر من عامين آخرين. منذ ذلك الحين ولحد اليوم وصلت طبعات الكتاب الى الطبعة السابعة وهو يحتل مركز الصدارة في الكتب المقررة في الجامعات العالمية لتلك المادة في أقسام الهندسة الكهربائية منذ ما يقرب من أربعة عقود، لكن لم تصدر ترجمة عربية لأي طبعة لاحقة لحد الآن ولم تستعمل الترجمة العربية إلاّ في العراق . ولو قدر لهذا الكتاب وحده أن تستمر ترجمات طبعاته وان تدرس مادته باللغة العربية في الجامعات العربية ، لأستعملت منه مئات الآلاف من النسخ على مستوى العالم العربي ولكان أساسًا مهمًا لتعريب التعليم في أقسام الهندسة الكهربائية. وهذا يدل على ضرورة العمل المؤسسي المستمر للترجمة بشكل عام وضرورة الترجمة الآلية بشكل خاص ، والحاجة إلى تكامل الجهود بين الجامعات العربية في مجال الترجمة.

والسبب الثاني في أهمية الترجمة الآلية هو أن عملية الترجمة للمترجمين من البشر عملية مملة وبطيئة والمترجم يحاول التغلب على مللله بتغيير اسلوبه تارة وبالراحة تارة وهو كطبيعة البشر ينام ويلهو ويمرض ويغيّر عمله ويتقاعد. كل ذلك يحدد عمل المترجمين ويجعلهم بضاعة نادرة في عصر العولمة. وإذا علمنا أن المترجم غالبًا ما يتقن لغة واحدة مع اللغة الأم التي يترجم منها أو إليها فيعني ذلك ندرة شديدة فيمن يتقن لغة كاللغة الكورية أو اليابانية مع العربية. وهذا يعطي أهمية اضافية للترجمة الآلية فالحاسوب يمكن أن يعمل ٢٤ ساعة في اليوم ولا يأخذ إجازة في نهاية الأسبوع ويمكن استبداله بما هو أفضل منه وتحسين إدائه وسرعته مع التقدم التقني الجاري اليوم. وهناك أسباب أخرى تقف غلى جانب الترجمة الآلية منها حاجة القطاع التجاري إلى ترجمة تعطي فكرة لا بأس بها عن المنتجات الصناعية والتجارية دون أن تكون على درجة عالية من الرصانة مع ازدياد حجم مثل هذه المعلومات المطلوب ترجمتها وتوسعها لتشمل لغات جديدة في عصر العولمة.

أن الحاسوب لديه قدرة جزئية في هذا الجانب لاشك . فكما هو معروف أن الحاسوب يعطيك ماتعطيه أولاً ، حيث إنه إنما يعطيك ما تسجله في ذاكرته أو تعلِّمه على فعله . ومن ثم كلما كان التحليل اللغوي وكانت القواعد اللغوية المقدمة للحاسوب والمخزونة في ذاكرته دقيقة أدى ذلك إلى ترجمات أقرب إلى الصواب.(١)

قابلية الحاسوب في خدمة الترجمة

التعادل أوالتقابل المعجمي والاصطلاحي يعتمد على ما نخزن في ذاكرة الحاسوب. وكذلك معرفة القواعد النحوية تعتمد على المجهود البشري ، كما تعتمد معرفة القواعد الأسلوبية على الجانب البشري ، أي على المعلومات والبرامج والإمكانات التي يوفرها مصممو البرنامج الحاسوبي ، فالحاسوب لديه القدرة الجيدة على التعامل مع هذه الأمور بسهولة كبيرة .

وهنا ينبغي أن نناقش إمكانية الآلة على القيام ببعض العمليات التي يقوم بها المترجم من البشر ، وسنعرض في ما يأتي لبعض منها:

١- معرفة الموضوع العام للمادة المطلوب ترجمتها: هناك الكثير من الأبحاث التي أدت إلى تطوير في برامج التعرف على الموضوع العام للمادة المراد ترجمتها أو أخذ فكرة عنها أو اختصارها. فمثلا يمكن للآلة استعمال دلائل تؤدي إلى معرفة إن كانت المادة سياسية أو اقتصادية أو علمية وإن كانت علمية هل هي في موضوع رياضي أو زراعي أو كيميائي مثلاً. إن تحديد موضوع المادة يساعد في تحديد المصطلحات المستعملة في ذلك العلم ومن ثم الوصول الى ترجمة أكثر دقة وتحديدًا. وبالطبع فإن دقة تعريف موضوع المقال أو البحث أو الفقرة المراد ترجمتها يعتمد على المصطلحات أو ما يسمى بالكلمات المفتاحية التي تحويها ، كما يعتمد على كيفية التعامل معها أو ما يسمى بمصطلح الحاسوب الخوارزميات المستعملة في تحديد الموضوع. وعلى ذلك فإن برنامجًا قد يكون أكفأ من برنامج آخر وفق الأسس التي اعتمدها واضعو ذلك البرنامج بالمقارنة مع البرنامج الآخر.

٢- استعمال المعاجم اللغوية بمقابلة كلمة مقابل كلمة ضمن الموضوع العام للمادة المطلوب ترجمتها. إن المعاجم قد تكون معاجم بلغة واحدة أو بأكثر من لغة. فالمعاجم بلغة واحدة تحدد معنى الكلمة في نطاق معناها بتلك اللغة وحقول استعمالاتها للوصول الى المفهوم المستعملة فيه تلك الكلمة. أما المعاجم بلغتين أو أكثر فهي تعطي ما يقابل الكلمة من كلمات في اللغة الثانية أو اللغات الأخرى. ويلاحظ هنا كثيرًا من التعقيد حيث هناك الكلمة يقابلها كلمة والكلمة مقابلها كلمتان أو أكثر وقد يكون المفهوم المتكون من أكثر من كلمة يقابله كلمة واحدة في اللغة الأخرى. وهناك الكلمة التي يصعب ايجاد مقابل لها في اللغة الأخرى. كل ذلك يحدد قابليات المعاجم المستعملة في خدمة الترجمة.

٣- تفكيك الكلمة الى مكوناتها فمثلاً للغة العربية يمكن تفكيك الكلمة الى مكوناتها من سوابق ولواحق كأحرف العطف وحروف الجر وعلامات الإعراب والضمائر المتصلة وغيرها وتطبيق قواعد الصرف على الكلمات ومعرفة وزنها الصرفي وفيما إذا كان مصدرًا أو إسم آلة أو غيرها. ولكن غياب التشكيل يجعل المهمة أصعب ووجود أكثر من احتمال للتفكيك أو التشكيل يجعل المهمة أعقد. ففي بعض الأحيان يصعب التمييز بين أن يكون الحرف في الكلمة حرفًا زائدًا كحرف جر أو حرف عطف أو أن يكون حرفًا أصليًا. فمثلاً كلمة "أهلك" هل هي فعل ماضٍ من الهلاك أم هي للمخاطب من الأهل؟.

٤- تحديد المفهوم العام لمعنى الجملة المراد ترجمتها ومحاولة إعادة صياغته لكي يكون مناسبًا للغة المراد الترجمة إليها. فقد أجريت أبحاث كثيرة لتحديد المفاهيم العامة في الحياة على مجموعات تتشعب منها مجموعات ثانوية وأخرى تتشعب عنها وفق تسلسل منطقي لتصنيف الأشياء الملموسة أو المعنوية. فمثلا التفاحة هي ثمرة والثمرة هي جزء من شجرة والشجرة نوع من النبات. وهكذا تصنف الأشياء تصنيفًا وصفيًا دقيقًا. وقد استعملت مثل هذه المفاهيم في الترجمة بين اللغات بحيث يكون هناك لغة وسيطة تترجم اليها المفاهيم من أية لغة ثم تترجم تلك المفاهيم من اللغة الوسيطة الى اللغة المراد الترجمة اليها. وقد توسعت هذه الأفكار وخاصة في جامعة الأمم المتحدة في اليابان. إلا أن نجاح هذه الطريقة لا يزال محدودًا لصعوبة تمثيل المفاهيم بشكل دقيق خاصة في الأمور المعنوية والأدبية.

٥- تركيب الجملة باللغة المراد الترجمة إليها وفق قواعد تلك اللغة . وهذه العملية تستعمل قوانين النحو والصرف والمعاجم لتركيب الجملة المراد تكوينها وفق المفهوم أو المعنى المستنتج من خطوات سابقة. وتخضع هذه الخطوات لاجتهادات اللغويين وواضعي البرامج الخاصة بذلك.

الحاسوب واستخداماته المختلفة في الترجمة

الحاسوب باستعمالاته اليوم وبالمستوى الذي وصلته عملية الترجمة الآلية لا يعوض عن المترجمين من البشر. وسوف تمضي بضعة عقود قبل أن تصل الترجمة الآلية إلى المستوى الذي يضاهي المترجمين الجيدين. لكن الحاسوب اليوم بإمكانه أن يقدم عونًا لا بأس به للمترجم . ومن هذه المعونة ما يأتي:

أ- الترجمة الآلية مع تحرير لاحق ، أي مراجعة بشرية بعد الترجمة الآلية. فالترجمة الآلية البدائية التي تعتمد على ترجمة معاني الكلمات وصياغة الترجمة بلغة ركيكة يمكن أن تكون بداية للمترجم لكي يقوم بإعادة صياغة الجمل وتنقيح المعاني ووضع الترجمة بشكل مقبول ومفهوم. وتجدر الإشارة أنه إذا كانت الترجمة الآلية سيئة جدًا فربما يكون من الأسهل على المترجم أن يقوم بالترجمة بنفسه من جديد بدل تنقيح نص سيئ جدًا.

ب- الترجمة مع التحرير السابق ، بمعنى أن الإنسان يحرر النص المراد ترجمته . مثلا يبسط الجمل المعقدة. والكلمات التي لها معانٍ كثيرة يحدد معناها المطلوب وهكذا ، أي أننا نعدل النص بحيث يستطيع أن "يفهمه" الحاسوب ، وتسمى هذه اللغة المقبولة للآلة Machine Acceptable Language . ويشبه ذلك التحاور مع الحاسوب بلغات البرمجة التي تتضمن كلمات محدودة بصيغ محددة لا يجوز التعدي لها.

جـ- هناك نوع ثالث يسمى بالترجمة التحاورية interactive وهي مثال للتعاون بين الحاسوب وبين المترجم البشري ، وذلك بأن يكون برنامج الترجمة ذا إمكانية حوارية بأن يعطي الترجمة جملة جملة ويتوقع من المترجم أن يوافق أو يعدل على بعض أجزائها لكي يصل إلى الترجمة المقبولة.

د- الترجمة البشرية بمعاونة الآلة ، أي أن الانسان يترجم والآلة تعاونه في هذه العملية . وهذا عكس الترجمة التحاورية . هنا نجد أن الانسان يترجم والآلة تبحث له في المعجم عن الكلمات وتعطيه معاني الكلمات كما تعطيه المرادفات من ذاكرتها .

هـ - الخدمات الحاسوبية الأخرى للترجمة مثل بنوك المصطلحات الآلية بشكل نصوص لمصطلحات متسلسلة أو البحث عن مصطلح فيها أو بإدخال مصطلح لمعرفة مرادفاته أو مقابله بلغة أخرى أو بالحصول على جميع المصطلحات في حقل معين إلى غير ذلك من الخدمات, (١)

تأريخ الترجمة الآلية

اول من استخدم الحاسوب في الترجمة بوضوح هو وارن ويفر عام ١٩٤٧. ومنذ عام ١٩٤٩ سارت بحوث الترجمة الالية في الولايات المتحدة قدمًا في جامعات كاليفورنيا ولوس انجلوس وتكساس وغيرها. وفي جورج تاون اجريت بنجاح اول ترجمة من اللغة الروسية الى الانكليزية عام ١٩٥٤. وفي عام ١٩٥٥ اجريت في الاتحاد السوفييتي اول تجربة في الترجمة الالية من الانكليزية الى الروسية في الرياضيات (على اساس قاموس يحوي ٢٣٠٠ كلمة) . وقد استمرت الأبحاث في هذه الفترة باستعمال ما يسمى الجيل الأول من برامج الترجمة الآلية حتى عام ١٩٦٦.(٦) وقد تبين خلال تلك الفترة حجم الصعوبات الهائلة التي تعترض الترجمة الآلية مما أدى إلى تباطؤ في أبحاث الترجمة الآلية وربما إهمال الموضوع حتى عام ١٩٧٥.

شهد الإهتمام بالترجمة الآلية في أوربا وكندا بين عامي ١٩٧٥ و ١٩٨٥ عودة إلى الأبحاث في هذا الحقل مدة عقد من الزمان تطورت فيه آلية استعمال الأنظمة الخبيرة وتقعيد معالجة اللغات الطبيعية واستحداث ما يمكن اعتباره الجيل الثاني من برامج الترجمة الآلية. وكانت نتيجة ذلك ظهور بعض برامج الترجمة الألية التجارية في الأسواق التي تستعمل الحواسيب المايكروية. لقد تطور في هذه الفترة البحث العلمي في معالجة اللغات الطبيعية وخاصة اللغات الأوربية واليابانية. وهذا التطور شمل البحوث المعجمية والنحو والصرف والدلالة.

في هذه الفترة تطورت كذلك أساليب الذكاء الإصطناعي الحديثة التي تستند إلى استعمال أنماط رياضية وحاسوبية تحاكي عمل الإنسان أو الكائنات الحية فأساليب الشبكات العصبية والنظم الخبيرة والمنطق المشوش والخوارزميات الجينية كلها تحوي أساليب رياضية تقلد مايجري في دماغ الإنسان أو تقلد أحياء خلقها الله وعلمها كيف تعيش في مجتمعات وكيف تتصرف كما هو الحال في محاكاة خلايا النمل وأنماط طيران الوزّ. وهذه الأساليب لا زالت في تطور مستمر إلى يومنا هذا.

وفي عام ١٩٨٩ بدأ عهد جديد في أسس الترجمة الآلية بالإستناد إلى المعلومات الإحصائية حينما قامت شركة أ.ب.م. بمشروع كارديد والذي استند إلى الترجمة بالأمثلة والترجمة المحدودة الموضوع وتعدد اللغات المترجم منها والمترجم إليها. وأعقب ذلك تزايد الإهتمام بالترجمة الآلية لحد لم يسبق له مثيل. وعلى ذلك يمكن اعتبار فترة التسعينيات من القرن العشرين فترة ظهور الجيل الثالث من برامج الترجمة الآلية المستندة إلى الذخيرة اللغوية Corpus Based MT والتي لا تزال في تطور حتى اليوم مع بعض التكامل مع الأساليب الأخرى للترجمة الألية.

واليوم تأتي الولايات المتحدة واليابان وروسيا والصين في طليعة البلدان التي توظف الترجمة الآلية لخدمة متطلباتها الاجتماعية والإقتصادية والعلمية والتقنية.

واقع الترجمة الآلية

هناك الآن ما يقرب من ١٠٠٠ برنامج ترجمة آلية (خاصة للغات الأوربية) في السوق رغم أن نوعيتها ليست جيدة بشكل عام لكن الطلب عليها عال جدًا. وقد زادت الأنترنيت من الحاجة إلى الترجمة الآلية وهي كذلك وسيلة سهلة لتسليم المادة المترجمة الى من يحتاجها. و سيشهد المستقبل تكاملاً بين عمل المترجمين من البشر والترجمة الآلية حيث تحتاج الترجمة الآلية الى مترجمين أكفاء لتطويرها ومتابعة عملها وإدخالها في مجالات لم تدخلها بعد. (٧)

إن استعمال أنظمة الترجمة الألية قليلة الدقة سيدفع الى الحاجة الى تحسين النوعية بشكل مطرد. و سينقلب التوتر السابق بين المترجمين من البشر ومؤيدي الترجمة الآلية الى تعاون بعد إدراك الجميع أن الترجمة الآلية لن تستطيع إلغاء دور المترجمين من البشر

تبلغ مبيعات برامج الترجمة الآلية في اليابان ملايين النسخ وبكلفة تبلغ عشرات الملايين من الدولارات. ويتوقع أن يتضاعف ذلك مئات المرات خلال العقدين القادمين مع التطور والتحسين المتوقع في الترجمة الآلية مستقبلاً.(٨)

أساليب الترجمة الآلية ـ محاكاة الآلة لعملية الترجمة البشرية

يمكن تصنيف برامج الترجمة الآلية بحسب مستوياتها التي تعكس مدى تعقيدها ومدى كفاءتها في الترجمة إلى الأصناف التالية بشكل تقريبي: (٩)

المستوى الأدنى

وهو باستبدال كلمة محل كلمة مكافئة لها. وهذه تحتاج إلى معجم ثنائي اللغة ضخم. وهنا ينبغي الأخذ بعين الإعتبار أن بعض الكلمات لا مقابل لها وبعضها تحتاج أكثر من كلمة وبعضها لها أكثر من مقابل وهكذا.

المستوى التالي الأعلى من المستوى الأدنى

وهو القيام بعلميات معالجة صرفية للوصول الى عبارات قياسية لغرض تقليص حجم المعجم المطلوب. وعلى ذلك فهناك حاجة الى التعامل مع الكلمات بمستوى مخفي بحيث توصف الكلمة بعدة مواصفات صرفية ونحوية. ويقع تحت هذا المستوى طريقة الترجمة بالأمثلة المباشرة حيث يكون هناك ذخيرة لغوية متوازية جملة جملة. وتكون العبارات المتوازية في الذخيرة مكونة من كلمتين فأكثر لكن المشكلة في هذه العبارات هي في مطابقة الجمل المراد ترجمتها مع ما هو موجود في الذخيرة.

المستوى المتوسط الأول

لكي يتم الحصول على الجملة المصدر بشكل صحيح يجب تكوين شجرة إعراب للجملة الأصلية ثم بعد ذلك إسقاط ذلك على اللغة المراد الترجمة إليها. وعلى ذلك فهناك حاجة إلى محلل نحوي بالإضافة إلى المعجم ثنائي اللغة.

المستوى المتوسط الثاني

هناك الكثير من الظواهر اللغوية التي لا يمكن نقلها من لغة إلى لغة بمجرد التحليل النحوي والصرفي. ففي العربية إذا قلنا " إن رأسي يؤلمني" فترجمتها بالأنكليزية هي " لدي صداع". وعلى هذا فمن الضروري فهم المعنى وتمثيل المعنى بشكل سليم لكي يمكن إيجاد المرادف له في اللغة الأخرى. وعلى هذا يجب تطوير المعجم لكي يحوي ترجمة لمثل هذه المعاني المتقابلة. ولهذا فإن معظم برامج الترجمة الآلية اليوم (مثل برنامج سيستران ولوغوس و يوروترا) تحوي محللاً صرفيًا ونحويًا ونوعًا من التمثيل الدلالي لمثل هذه الحالات.

المستوى الأعلى

في هذا المستوى يجب أخذ الأساليب البلاغية العميقة في اللغة المصدر واللغة المترجم إليها ولا زال البحث في هذا المستوى يأخذ أبعادًا مختلفة ويحتاج إلى دراسات لغوية وتمثيل حاسوبي عميق ويمثل قصورًا واضحًا في برامج الترجمة الآلية المتوفرة اليوم.

اللغة الوسيطة

هناك اتجاه باستعمال لغة وسيطة بين اللغات. وقد لاقت فكرة اللغة العالمية الوسيطة صعوبات في أن تحل كل مشاكل الترجمة الآلية. وقد اقترح أن تكون اللغة الأنكليزية بوضعها الحالي هي اللغة الوسيطة بين اللغات العالمية جميعًا. وقد يكون من المناسب أن تكون هي اللغة الوسيطة بين اللغات الأوربية الحديثة كوريث للغة اللاتينية التي انبثقت منها معظم اللغات الأوربية. ولكن أليست اللغة العربية جديرة بأن تكون هي اللغة الوسيطة بين اللغات التي يتكلم بها المسلمون اليوم كالتركية والفارسية والأوردية والسواحيلية والملاوية والبنغالية التي قد اشتقت كثير من كلماتها من العربية؟ هذا سؤال على العرب اليوم أن يجيبوا عليه.

وفي كل الأحوال على برنامج الترجمة الآلية أن يقوم بعمليتين رئيسيتين:

(أ) النقل ، أي عملية إيجاد المقابلات المعجمية والنحوية والأسلوبية لأجزاء النص المترجم ، مثل إيجاد الكلمات

والتعبيرات المقابلة في المعني والوظيفة للكلمات والتعبيرات الواردة في النص الأصل ، وكذلك إيجاد التراكيب النحوية المقابلة للتراكيب الواردة في النص الأصل

(ب) أما العملية الفرعية الثانية فهي التأليف أو التوليف ، أي صياغية الجمل الناتجة من عملية النقل السابقة صياغية صحيحة صرفيًا ونحويًا وأسلوبيًا، مثل صوغ (رجل +جمع) في صورة (رجال) و(مسلم +جمع+حالة النصب أو الجر) في صورة (مسلمين) ، وكذلك وضع الصفة في العربية بعد الموصوف ومراعاة قواعد المطابقة اللازمة.(١)

الترجمة الآلية المستندة إلى الإحصاء

تستند هذه الطريقة الحديثة الى جمع أكبر ما يمكن من ذخيرة لغوية (corpus) وإعمال أكبر ما يمكن من جهد إحصائي عليها لكي تهيأ للاستخدام في الترجمة الآلية. وهذه الذخيرة المترجمة هي بالأساس مترجمة من قبل مترجمين من البشر فهي تستخلص خبرات البشر للإفادة منها في الترجمة الآلية.

الفلسفة وراء هذه الطريقة هي أن الذخيرة اللغوية إذا كانت كبيرة الحجم بما فيه الكفاية فهي تجمع بين دفاتها معظم الكلمات الشائعة في اللغة ومعظم التعابير اللغوية ومعظم التراكيب النحوية والصرفية فيها. وعلى ذلك فإن أي عملية إحصائية على هذه الذخيرة تعطي نتائج قريبة من واقع اللغة الفعلي.

ويحق لنا أن نتساءل عن مدى صحة هذه الفرضية نظرًا لأهميتها واعتماد نظرية الترجمة الألية عليها بشكل كبير.

وللإجابة على هذا التساؤل يجب أن نحدد أولاً عن أية ذخيرة لغوية نتحدث؟. ومن أين جمعت؟ وكيف جمعت؟ وأي المواضيع اللغوية تشمل؟ ولأي نوع من الترجمة ستستعمل؟ وهل هي بلغة واحدة أم متعددة اللغات؟ وهل هي جاهزة للتقابل اللغوي؟

ما يحدث اليوم في اللغات العالمية الأخرى غير العربية هو افتراض أن اللغة المتوفر بها كتابات كافية كالصحف والمجلات والكتب المطبوعة والمقالات المنشورة على الأنترنيت والأعلانات ونشرات الشركات التجارية تمثل اللغة. ومن ثم استناد الترجمة الآلية لمثل هذه الذخيرة يمثل هذه اللغة أو تلك. ويعتبر حجم الذخيرة الذي يبلغ مئات الملايين من الحروف (مئات الميكابايت) مقدارًا يمثل الحدود الدنيا لحجم الذخيرة التي يمكن الإعتماد عليها للحصول على ترجمة معقولة ومقبولة. وقد سبق وأن ذكرنا الفقر الشديد الذي تعاني منه الترجمة من العربية وإليها ، وعلى ذلك فالذخيرة اللغوية الممكن توفرها باللغة العربية لا تزال محدودة.

ويقع تحت هذا الأسلوب من الترجمة عدة طرائق منها : الترجمة الإحصائية المباشرة والترجمة بالأمثلة والترجمة بالتجاور، ففي الأسلوب المباشر يكون هناك معجم عبارات متقابل بين اللغات المراد الترجمة منها واليها . وفي الأسلوب غير المباشر يتم إجراء تحويلات لغوية متعددة للإفادة من الذخيرة كالتحويلات الصرفية والنحوية والدلالية واستعمال الإحصاء في الأنظمة الخبيرة.

هذه الطريقة تحتاج إلى أعراب الجملة نحويًا بلغتها الأصلية ثم تطبيق جملة من التحويلات على شجرة الإعراب المستحصلة أعلاه وذلك بإعادة تسلسل السلسلة الظاهرية من جهة اللغة الأصلية وذلك للوصول الى تسلسل أفضل لكلمات الجملة بحيث تكون أقرب الى اللغة المراد الترجمة إليها وهذه العملية تطبق على الذخيرة المتعملة للتدريب كما تطبق على النصوص المستعملة في الترجمة

كيف تجرى العملية الإحصائية على الذخيرة:

يقوم الحاسوب أولا باستخراج الكلمات غير المكررة في الذخيرة ليعمل منها قائمة بكلمات الذخيرة. ثم بعد ذلك يقوم بعمل قائمة بكل كلمتين متعاقبتين فيها ليكوّن من ذلك قائمة بالعبارات المكونة من كلمتين فيحدد مرات ورود كل منها، ثم يقوم بعمل قائمة بثلاث كلمات متعاقبة ويعمل منها قائمة تحوي عدد مرات ورود هذه العبارات المكونة من ثلاث كلمات وهكذا يزيد في أعداد الكلمات إلى أن لا يبقى هناك كلمات متعاقبة مكررة أخرى أو يتوقف عند عدد محدد من الكلمات.

هذه العملية الإحصائية هي واحدة من العمليات الإحصائية المهمة التي تستند إليها كثير من أبحاث المعالجة الآلية للغات الطبيعية اليوم. لكن الترجمة الآلية تحتاج إلى نصوص متوفرة بلغتين أو أكثر. فالذخيرة اللغوية يجب أن تكون متوفرة بلغتين ويجب أن تكون مصفوفة بشكل متقابل.

تستند هذه الطريقة إلى عمل جداول بالعبارات المكررة بين اللغتين والمكونة من كلمتين أو أكثر. فترجمة عبارة معينة بأكثر من موقع واحد في الذخيرة بالترجمة نفسها وربما من قبل مترجمين مختلفين يزود برنامج الترجمة الآلية بترجمة مفضلة لتلك العبارة. أما إذا اختلفت الترجمة بين مصدر من مصادر الذخيرة وآخر أو في المصدر نفسه فتتبع الأساليب الإحصائية باختيار الترجمة الأكثر تكرارًا أو التي يقع الموضوع الذي وردت فيه ضمن موضوع النص المراد ترجمته. هذا وإن استخلاص العبارات المحتوية على كلمتين أو أكثر يتم ببرامج حاسوبية متوفرة اليوم.

هذه الطريقة بأسلوبها البسيط تعطي ترجمة غير دقيقة بالتأكيد ، لكن إلحاق معلومات نحوية وصرفية مع المعجم المستخلص من الذخيرة سوف يزيد من دقة الترجمة. وقد استعملت هذه الطريقة على ترجمة محاضر البرلمان الكندي بين اللغتين الفرنسية والأنكليزية فوجدت ناجحة لحد كبير وذلك بسبب أن موضوع التحاور في أروقة البرلمان ذو نمط معين وبأساليب متعارف عليها بين المتحاورين. ويعني ذلك أن هناك اسلوبان للإفادة من الذخيرة اللغوية وهما الأسلوب المباشر والأسلوب غير المباشر.

الترجمة بالأمثلة

تستند هذه الطريقة إلى اجراء ابحاث على الذخيرة اللغوية لاستخلاص الأمثلة والعبارات الشائعة المتقابلة بين اللغتين المراد الترجمة بينهما. كما أن استخلاص قوالب نحوية بين اللغتين وتعويض الكلمات المقابلة بين اللغتين هو أحد الأساليب التي تقع ضمن هذا الأسلوب. هذه الطريقة عادة فعالة للترجمة بين اللغات التي تعود للعائلة نفسها كالأنكليزية والفرنسية أو العربية والسواحيلية بسبب سهولة تكوين قوالب متقابلة بين اللغتين العائدتين للعائلة نفسها لكن الأمر يزداد صعوبة بين اللغات التي تعود إلى عوائل مختلفة

لناخذ العبارات القرآنية المتشابهة التركيب اللغوي مع اختلاف كلمة فيها (٤)

فـ + اعلموا أن الله + ... ٣ آيات
و + اعلموا أن الله + ... ١٢ آية

والله (بصير ، خبير ، عليم) بما (تعملون ، يفعلون ، يصنعون ، يعملون)

والله بما (تـ ، يـ) ـعملون (محيط ، بصير ، خبير ، عليم)

كأنهم أعجاز نخل (خاوية ، منقعر)

مثل هذه العبارات القرآنية يمكن أن تشكل قوالب للترجمة بالأمثلة بحيث إذا تغيرت كلمة فيها فإن القالب يكون نفسه باستبدال ترجمة الكلمة الواحدة فقط.

الترجمة بالتجاور Connectionist

تستند هذه الطريقة إلى استعمال وسائل الذكاء الإصطناعي كالشبكات العصبية وتحليل الجملة الى شبكة تستعمل كمعلومات لما يرد من جمل لاحقة بحيث يزداد البرنامج خبرة كلما استعمل مرة بعد مرة فهو يخزن خبراته لكي يستعملها مستقبلاً. فإذا ما قام المستخدم بتعديل معنى معين سبق وأن أخطأ فيه البرنامج فإنه يتعلم من أخطائه ولا يكرر الخطأ بل يسلك الأسلوب الجديد بدل ذلك. كما سيأتي ذكره في برنامج جوجل.

المشاكل التي تعاني منها الترجمة الإحصائية

لا تزال الترجمة الإحصائية تعاني من كثير من الصعوبات . من هذه المشاكل أن الدقة في الترجمة الإحصائية تعتمد على نوع ودقة الذخيرة المستعملة وعلى البرمجيات المستعملة فيها.

كما إن تغير اللغة والمفردات المستعملة مع مضي الزمن وضخامة الذخيرة المطلوبة يجعل هناك ضرورة لتحديد الزمن التأريخي للنص . فكلمة "السيارة" حين ترد في نص مكتوب قبل قرون تعني غير " السيارة" المكتوبة في نص مستعمل اليوم. كما أن هناك مشكلة في تحديد الأسماء فكلمة "أحمد" في جملة " أحمد الله على نعمائه" هي ليست إسم علم . ومن الصعوبة بمكان على الآلة تحديد هل الكلمة ذات معنى أم إسم علم. وتحتاج إلى الكثير من الضوابط البرمجية لكي تميز الآلة بين الكلمة كإسم علم أو غيره.

حل الغموض بواسطة المعالجة المتوازية

تتلخص طريقة الترجمة الآلية بالتوازي بأن تستعمل أكثر من طريقة واحدة لترجمة نص معين في وقت واحد، ثم بعد ذلك يجري تقييم هذه الترجمات والمفاضلة بينها لاختيار أفضلها وربما يجري التزاوج بين بعضها للوصول إلى ترجمة تجمع محاسن أكثر من طريقة واحدة بوقت واحد.

بعد إجراء هذه التحليلات نصل الى احتمالات متعددة للترجمة الوسيطة يجري تقييمها واسقاط ما هو أقل احتمالا منها ومن ثم الوصول الى الترجمة الأفضل بحيث يعاد تركيب الجملة في اللغة المراد الترجمة اليها بخطوات معاكسة.

الترجمة الآلية واللغة العربية

تتشابه المشاكل التي تعاني منها الترجمة من اللغة العربية وإليها مع المشاكل التي تعاني منها اللغات الأخرى ، فإذا ما استثنينا الشعر والنثر الأدبي الراقي والمحتوي على دلالات بلاغية عالية فإن هذا ينطبق على اللغة العربية أيضًا. وفيما يأتي بعض خصائص اللغة العربية ذات العلاقة بالترجمة ، فليس المجال هنا للكلام عن خصائص اللغة العربية بشكل عام ولكن هناك بعض الخصائص التي تؤثر على التقدم في حقل الترجمة الآلية بشكل خاص وهو موضوعنا هنا. فمن الخصائص المهمة للغة العربية المستعملة بكثرة اليوم هو غياب التشكيل. ورغم أن

هناك محاولات لكتابة برامج للتشكيل الآلي للنص غير المشكول ، إلا أن أقصى دقة قد تصل إليها مثل هذه البرامج اليوم لا يتعدى نسبة ٩٥% . ويزيد الخطأ عن ٥% بكثير بالنسبة لتشكيل أواخر الكلمات. إن غياب التشكيل يجعل القارئ يفترض من عنده تشكيلاً افتراضيًا بما يملكه من ثروة لغوية. وكذلك يقوم المترجم. أما الآلة فيجب تعليمها ذلك. (١٠)

كما أن ندرة استعمال علامات الوقف والفواصل في النصوص العربية يضيف تعقيدًا آخر للنص العربي.

أما من ناحية الصرف والنحو وطريقة الكتابة والعمليات التي تجري على الكلمة من دمج للسوابق واللواحق وإدغام واقلاب وغيرها فهو ما تختص به العربية. ولكل لغة خصوصياتها.

وكثيرًا ما تكون الجملة العربية طويلة ، حيث يبلغ متوسط عدد كلمات الجملة العربية ما يزيد على ٣٥ كلمة. (١١) وهذا يؤدي الى بطئ المعالجة الحاسوبية للجملة العربية. ويمكن التغلب على هذا الطول بتقسيم الجملة الى عبارات ....مثلا المضاف والمضاف اليه وأشباه الجمل من جار ومجرور ...

لكن هناك بعدًا آخر في اللغة العربية ينبغي الإنتباه اليه هو البعد التأريخي. فالنصوص الأنكليزية مثلاً التي تتضمنها الذخيرة لا يتعدى تأريخها القرن أو القرنين من الزمان وهي مع هذا قد يلاحظ فيها التغير بين كتابات القرن الماضي والقرن الحالي. وهذا الفرق قد يوازي الفرق بين اللغة العربية في صدر الإسلام (والتي كتبت بها معظم المصادر الدينية والتأريخية ولم تتغير الا قليلا) واللغة العربية اليوم.

<u>المعالجة الإبتدائية للغة العربية قبل الترجمة الإحصائية</u>

تتضمن العمليات االتمهيدية للغة العربية قبل الترجمة الآلية إجراء عمليات عديدة منها العمليات الآتية وهذه على سبيل اللمثال لا الحصر . (١٢)

١- المعالجة الكتابية للنص فيما يتعلق بالهمزات والهاء الآخرية والتاء المربوطة والتشكيل.
٢- تجزئة الكلمة إلى سوابق ولواحق وتحديد أجزاء كل منها إن وجدت كألف لام التعريف وحروف العطف وحروف الجر المتصلة والضمائر وعلامات الإعراب وتحديد جذر جذع الكلمة ووزنها الصرفي.
٣- الرجوع إلى قواعد تغيير كتابة الكلمة بعد إلحاق بعض الضمائر المتصلة مثل كلمة مكتبة – مكتبتهم وأعلى – أعلاه ..
٤- تحديد فيما إذا كان العدد المتضمن ضمن الكلمة مفردًا أو مثنى أو جمع وتحديد علامة ذلك كالألف والنون أو الواو والنون وغيرها.
٥- حل بعض المشاكل المتوقعة لوجود أكثر من احتمال في تحديد الكلمة المراد ترجمة معناها مثل الكلمات التي تعطي أكثر من معنى باختلاف التشكيل كأن يكون الفعل مبنيًا للمجهول أو للمعلوم أو أن يكون اسمًا أو فعلاً مثل كَتَب كُتِب كُتُب

هذه بعض العلميات التي يجب القيام بها أو ببعضها بحسب البرامج التي تحتاجها الترجمة الآلية والتي يجب القيام بها سواء على الذخيرة اللغوية أو على النص المراد ترجمته من العربية أو النص المراد توليده باللغة العربية إن

كان مترجمًا من لغة أخرى. ولنأخذ العبارات القرآنية الآتية كمثال على العبارات المكونة من كلمتين والواردة في القرآن الكريم بكل أشكالها:

وَيُقِيمُونَ الصَّلاَةَ ، يُقِيمُونَ الصَّلاَةَ ، لِيُقِيمُوا الصَّلاَةَ ، وَيُقِيمُوا الصَّلاَةَ ، يُقِيمُوا الصَّلاَةَ ، أَقَامُوا الصَّلاَةَ ، وَأَقَامَ الصَّلاَةَ ، وَأَقَامُوا الصَّلاَةَ ، أَقَمْتُمُ الصَّلاَةَ ، أَقِمِ الصَّلاَةَ ، أَقِيمُوا الصَّلاَةَ ، فَأَقِيمُوا الصَّلاَةَ ، وَأَقِمِ الصَّلاَةَ ، وَأَقَمْنَ الصَّلاَةَ ، وَأَقِيمُوا الصَّلاَةَ ، وَإِقَامَ الصَّلاَةِ ، وَإِقَامِ الصَّلاَةِ ، قَائِمٌ يُصَلِّي ، مُقِيمَ الصَّلاَةِ ، وَالْمُقِيمِي الصَّلاَةِ ، وَالْمُقِيمِينَ الصَّلاَةَ ، فَأَقَمْتَ لَهُمُ الصَّلاَةِ

فهذه العبارات كل منها يحتوي على كلمة تتعلق بإقام الصلاة مع بعض الإضافات من سوابق أو لواحق أو تغيير في الكلمة الأولى بين الفعل والإسم. فإذا ما حددت الإضافة من ناحية مفعوليتها أمكن ترجمة ذلك إلى اللغة الثانية . فإذا عرفنا من ترجمة معينة ما معنى "إقام الصلاة" بتلك اللغة , أمكن ترجمة كل العبارات السابقة بالإستناد إلى ذلك المعنى بعد الحاقه وتغيير تركيبته وفق قواعد محددة تترجم الإضافات الموجودة في اللغة الأصلية. ويلاحظ أن بين العبارات السابقة العبارة الأخيرة المكونة من ثلاث كلمات : **فَأَقَمْتَ لَهُمُ الصَّلاَةِ** حيث أن الكلمتين المعنيتين دخل بينهما جار ومجرور "لهم" ومثل هذه العبارات يجب البحث عنها أيضا لتلحق بمثل هذه المجاميع من العبارات: وكمثال على ثلاث كلمات بينها ضمائر أو ما يقابل تلك الضمائر فتصبح أربع كلمات :

أَنفِقُوا مِمَّا رَزَقْنَاكُم ، أَنفِقُوا مِمَّا رَزَقَكُمْ ، وَأَنفَقُوا مِمَّا رَزَقْنَاهُمْ ، وَأَنفَقُوا مِمَّا رَزَقَهُمْ ، وَيُنفِقُوا مِمَّا رَزَقْنَاهُمْ ، وَأَنفَقُوا مِمَّا رَزَقَهُمُ اللَّهُ ، أَنفِقُوا مِمَّا رَزَقَكُمُ اللَّهُ ، وَمِمَّا رَزَقْنَاهُمْ يُنفِقُونَ

ويلاحظ هنا تغير تسلسل الكلمات بين عبارة وأخرى. وعلى ذلك فإن الترجمة يجب أن تأخذ بعين الإعتبار ما هو متعارف عليه باللغة المترجم إليها بعد معرفة مفهوم النص وهو "الإنفاق من رزق الله". وكذلك العكس عند الترجمة إلى اللغة العربية.

وصف الكلمة وفق مشروع مداد البيان(٥)

من الأمثلة السابقة يتبين أن المعالجة اللغوية للنصوص الأصلية قبل أن تكون جاهزة للترجمة الآلية ضرورية لتمام الإفادة من الذخيرة اللغوية والتي يجب أن تكون بحجم كافٍ لكي تغطي اللغة أو معظم الأساليب الواردة فيها على الأقل. وقد احتوى مشروع مداد البيان الخاص بوصف كلمات القرآن الكريم في قواعد بيانات تفصيلية على معالجات للكلمات القرآنية يمكن أن يخدم اللغة العربية بشكل عام وكمثال على بعض المعالجات التي ينبغي القيام بها في اللغة العربية عند تحليل الأمثلة أو النص المراد ترجمته أو عند توليد نص مترجم المعالجات الآتية ، نورد ما يأتي :

١- الإدغام مثل **مُضَارّ وَازَّيَّنَتْ**
٢.الإبدال. مثل **وَاصْطَبِرْ**
٣. الإعلال بالقلب. مثل: **قِيلَ أو** بالحذف مثل: **قُلْتُ** أو بالنّقل.مثل: **وَيَقُولُونَ** أو بالحذف الصوتي مثل: **نَبْغِ**
٤. حذف الهمزة.: مثل: **كُلُوا** و تسهيل الهمزة مثل: **بَادِيَ**

٥- حذف ياء المنقوص مثل: **بَاغٍ** وَلا **عَادٍ**
٦. حذف النّون للإضافة مثل: **وَالْمُقِيمِي**
٧. حذف حرف العلّة للجزم: وَلا **تَمْشِ**
٨. حذف نون الأفعال الخمسة: وَلا **تَقُولُوا**

وعلى ذلك فإن التدريب الذي يجب ان يتم يحتاج كمية كبيرة من الأمثلة التي تحوي مثل هذه الحالات لكي يكون كافيًا للقيام بعملية الترجمة بشكل مقبول.

الذخيرة اللغوية العربية

تعتبر الذخيرة اللغوية لأية لغة مرتكزًا أساسيًا اليوم للباحثين في اللغات الطبيعية. فالخزن على الحواسيب جعل جمع وتصنيف وتحليل الذخيرة اللغوية ميسورًا. وهذه الذخيرة تمكن الباحثين من إطلاق الوصف على خصائص اللغة وعلى النحو فيها ودراستها تأريخيًا وتغيرها مع الزمن.(١٣)

كانت أولى محاولات تكوين ذخيرة لغوية عربية صغيرة في عام ١٩٩٤. ومن ثم بدأت محاولات متعاقبة لجمع ذخيرة لغوية عربية ذات شأن. ومن هذه المحاولات جمع نصوص جريدة الحياة اللندنية ما بين ١٩٩٤ و ٢٠٠٠ والتي احتوت على ما يقرب من ٧٦ مليون كلمة منها ٦٦٦ ألف كلمة غير مكررة. كما حدثت محاولات عديدة أخرى بتكوين ذخيرة لغوية من مستندات الأمم المتحدة متعددة اللغات منها محاولة لعدد من الوثائق بلغ ٣٨ ألف وثيقة تحوي ما يزيد على ٥٠ مليون كلمة.

يزداد استعمال اللغة العربية على شبكة الأنترنيت بسرعة هائلة. ففي الوقت الذي ازداد استعمال اللغة الإنكليزية بين عامي ٢٠٠٠ و ٢٠٠٨ بنسبة ٢٠٠% ازداد استعمال اللغة العربية خلال الفترة نفسها بنسبة ٢٠٠٠% بينما كانت الزيادة في استعمال اللغة الصينية بنسبة ٦٠٠% خلال الفترة نفسها.

يبلغ عدد مستخدمي الأنترنيت باللغة العربية اليوم حوالي ٦٠ مليون مستخدم أي بنسبة ٤,٢% من عدد المستخدمين في العالم وهو يقرب من نسبة عدد السكان العرب بالنسبة لمجمل سكان العالم والبالغ حوالي ٤,٧% (١٤) ويعني ذلك أن نمو استخدام الأنترنيت في العالم العربي قد وصل ما يقرب من المعدل العالمي وهو بازدياد. ولكن هذا بالطبع لا يعكس مستوى المحتويات العربية حيث لا تزال أقل من ١% من محتوى الأنترنيت حسب توزيع اللغات.

بعد اعتماد الرمز الدولي الموحد المعروف باليونيكود في نهاية القرن الماضي ، احتلت الحروف العربية مواقع محددة في جدول الرموز مما يعني أن أي وثيقة على الأنترنيت تستعمل هذه الرموز يمكن مباشرة معرفة فيما إذا كانت هذه الوثيقة تستعمل الحروف العربية أم لا. ومن ثم يمكن معرفة لغتها غالبًا.

إن الذخيرة اللغوية يمكن أن تكون مصدرًا مهمًا للمعجم التأريخي وذلك بإدخال معلومات عن تأريخ كتابة النص الذي تتضمنه الذخيرة اللغوية كما ذكرنا أعلاه في معنى كلمة السيارة مثلاً. فترجمة كلمة كهذه يجب أن تأخذ بعين الإعتبار تأريخ النص المستعمل في الذخيرة اللغوية وما يقابها من معنى في النص المطلوب ترجمته. لأن اختلاط الأمر قد يؤدي إلى ترجمات غير صحيحة كما هو متوقع.

باستعمال ذخيرة محدودة يمكن توسيع تطبيقها على امثلة اوسع بكثير منها بتحديد شروط معينة لما يمكن ان يقاس على هذه الذخيرة واضافة كلمات مركبة بدل الكلمات الواردة في الذخيرة وجداول احتمالية لما يمكن ان يضاف في المستقبل. فمثلا إذا وضعت ذخيرة تشمل كلمات القرآن الكريم فإنه ليس من الصعوبة بمكان تطوير الذخيرة إلى الحديث النبوي مثلاً.

خطوات تكوين ذخيرة لغوية (١٥)

١- الحصول على نص مكتوب متوفر إلكترونيًا وليس عليه حقوق ملكية فكرية. وتعتبر شبكة الأنترنيت أيسر المصادر المتوفرة. يمكن الوصول الى الصفحة الموضوعة على الأنترنيت عن طريق التنقيب باستعمال ما يسمى بالبرامج العنكبوتية بمساعدة رموز HTML . وبعد تنقية النصوص من هذه الرموز يتم الوصول الى النصوص الخام. و قد تحتاج هذه التنقية في أغلب الأحيان إلى تدخل يدوي . ومن ثم يتم توضع رموز خاصة على هذه النصوص للإشارة الى مصادرها ومعلومات أخرى عنها. وحيث أن طريقة ترميز النصوص العربية مختلفة بين وثيقة وأخرى بحسب مواصفات الترميز العربية المختلفة لذلك ينبغي أن توحد طريقة ترميزها ليسهل التعامل معها. ويعتبر اليوم الترميز الدولي الموحد (اليونيكود) رغم نقائصه أفضل ترميز لأغراض توحيد النصوص.
بعد إجراء عمليات عديدة لتنقية النصوص من بعض الرموز كالمدة والتشكيل (إن وجد نظرًا لأن الغالبية العظمى من النصوص هي غير مشكولة) وأنواع الهمزة والتاء المربوطة والهاء الآخرية وغيرها مما يكثر اللبس فيها. ويلاحظ أن مثل هذا التوحيد يزيد من الغموض واحتمالات الخطأ كما يخفي في طياته بعض المؤشرات حول مصادر النصوص وأماكن تداولها وغير ذلك من المعلومات.
الذخيرة ما هي إلا خزان ضخم للغة المستعملة. والبرامج الخاصة بالذخيرة هي وسائل لإعادة ترتيب هذا الخزان لكي يكون بالإمكان استخلاص الملاحظات عن اللغة من خلالها. وتشير الأبحاث إلى أن الشبكة العنكبوتية اليوم مؤهلة تمامًا للحصول على ذخيرة لغوية مناسبة للغة العربية من خلال نصوص الصحف اليومية والصفحات التجارية والشخصية والكتب القديمة والحديثة المتوفرة على الشبكة. إلا أن النصوص المترجمة بلغة أخرى توازي اللغة العربية ، فيما عدا وثائق الأمم المتحدة ، لا تزال قليلة لا تكفي لتكوين أساس لذخيرة لغوية متعددة اللغات يمكن الإعتماد عليها بكفاءة.
٢- عملية التوافق بين النصوص في الذخيرة المكونة لأكثر من لغة واحدة. يجب مقابلة النصوص المترجمة مع بعضها البعض بشكل دقيق فقرة فقرة. وإذا أمكن وضعها جملة جملة أو عبارة عبارة فذلك أفضل بالطبع ولكن في العادة لا يمكن ذلك تمامًا. وهنا يجب حل بعض جوانب الغموض. والجدول الآتي يبين مثالاً لذلك:

Narrated 'Umar bin Al-Khattab: I heard Allah's Apostle peace be upon him saying,	عن عمر بن الخطاب قال: سمعت رسول الله صلى الله عليه وسلم يقول:
The reward of deeds depends upon the intentions	إنما الأعمال بالنيات،
and every person will get the reward according to what he has intended.	وإنما لكل امرىء ما نوى،
So whoever emigrated for worldly benefits or for a woman to marry, his emigration was for what he emigrated for.	فمن كانت هجرته إلى دنيا يصيبها، أو إلى امرأة ينكحها، فهجرته إلى ما هاجر إليه

٣- يلاحظ في أغلب الأحيان أن الكلمات ومعانيها المقابلة لها بشكل منفرد هي الغالب أي أن الغالبية العظمى من كلمات الذخيرة لها مقابل واحدة أي كلمة مقابل كلمة. وكذلك عبارة مقابل عبارة حتى ولو كان هناك في بعض الأحيان قلب لتسلسل الكلمات ضمن العبارة الواحدة.

٤- بعد ذلك يمكن جمع الإحصائيات عن تكرار الكلمات والعبارات والتركيبات الهيكلية في النصوص وهناك برامج متخصصة لاستخلاص المعلومات الاحصائية عن الذخيرة. ومن هذه الإحصائيات يمكن استخلاص معاني الكلمات والعبارات بحسب أعلى تكرار لها وحل بعض مشاكل الغموض الذي يحدث.

٥- يلاحظ في النصوص وجود نقاط ارتكاز لمقابلة التراجم مع بعضها البعض مثل وجود الأرقام في النص والتواريخ وأسماء الأعلام وعناوين الفقرات. ومثل نقاط الإرتكاز هذه بالإضافة إلى النقاط والفوارز بين الجمل يمكن استعمالها لترتيب الذخيرة بشكل متوازي . وينتج عن ذلك ثلاث أنواع من التقابل بين النصوص:

أ – تقابل قوي حينما تحوي النصوص وترجمتها على عدد متساوي من الكلمات المتقابلة وهي حالات نادرة .

ب- تقابل متقارب حينما يكون عدد الكلمات نفسه لكن هناك اختلاف في تسلسل الكلمات.

ج – تقابل ضعيف حينما يكون تسلسل الكلمات وعددها مختلف ولكن المعنى متوافق وفق المعجم، وعلى العموم فإن:

١- هناك علاقة معنى بين فصلين مترجمين متقابلين
٢- هناك علاقة معنى بين كل فقرتين مترجمتين متقابلتين
٣- هناك علاقة معنى بين كل جملتين مترجمتين متقابلتين
٤- هناك علاقة معنى بين كل كلمتين مترجمتين متقابلتين إذا كانت إحدى معاني الكلمات واردة في المعجم.
٥- يعتبر ورود الأرقام والتواريخ نقاط ارتكاز للتقابل بين النصوص المترجمة
٦- تسعمل نظرية الإحتمالات لحل بعض الغموض الذي يحيط بالتقابل بين النصوص

الأبحاث الغربية مؤخرًا في الترجمة من اللغة العربية وإليها

بعد الحملة الشرسة على الأمة بحجة أحداث الحادي عشر من ايلول برزت الحاجة لدي الدوائر العسكرية والمخابراتية والأعلامية الغربية للترجمة من اللغة العربية خاصة. فبعد الإحتلال الأمريكي للعراق جُهز بعض الجنود بأجهزة تشبه الهاتف النقال بحيث يتكلم الجندي بكلمات او عبارات باللغة الأنكليزية فتنطق ترجمتها بالعربية لكي يسمعها العراقي وبالعكس. كانت الكلمات والجمل محدودة وقد سجلت باللهجة العامية العراقية. إلا أن كل ذلك توقف خلال اسابيع عند بدء المقاومة العراقية وتوجس جنود الأحتلال بعد ذلك من التحاور مع السكان المدنيين خوفًا على حياتهم من أن يكون من يقترب منهم هو من أفراد المقاومة. (١٦)

وقد شهدت كثير من الجامعات الأمريكية والأوربية جهودًا حثيثة للبحث عن معالجة اللغة العربية حاسوبيًا والترجمة منها وإليها. وقد شملت هذه الأبحاث البحث في النحو والصرف والدلالة والترجمة وجمع الذخيرة اللغوية وكيفية تكوين المعجم المحوسب والإستفسار باللغة العربية وتكوين الخلاصات وغير ذلك من الأبحاث المتعلقة باللغة العربية.

المعجم العربي المحوسب

يكاد أن يتفق الحاسوبيون العرب المهتمون بمعالجة اللغة العربية على أن من أهم ما تحتاجه اللغة العربية اليوم لدخولها عالم تقنية المعلومات بفاعلية هو تكوين معجم عربي محوسب يتضمن كل مفردات اللغة العربية بشكل يسهّل التعامل مع مفرداتها في كافة التطبيقات ذات العلاقة. إن السبب في عدم تكوّن مثل هذا المعجم لحد الآن هو عدم وجود جهة علمية ذات قدرة مالية على مستوى العالم العربي تأخذ على عاتقها القيام بمثل هذا المشروع الضخم والبالغ الأهمية.

سنحاول هنا أن نبين أهمية المعجم والخطوة الأولى في تكوينه وهي وضع مواصفة له بحيث يبني عليها من يأتي من بعد.

ما المقصود بالمعجم العربي المحوسب

تتوفر للغة العربية قواميس تعطي معاني كلماتها للغات أخرى. كما تتوفر قواميس عربية عربية تعطي معاني الكلمات بكلمات أخرى. ويوضع في هذه القواميس عادة تصريف الكلمة وبعض المدخلات الأخرى. وقد تعددت المعاجم العربية وتنوعت خلال العصور السالفة ولكن القصد منها في كل الأحوال كان واحدًا وهو حراسة القرآن من أن يقتحمه خطأ في النطق أو الفهم، وحراسة العربية من أن يقتحم حرمها دخيل لا ترضى عنه العربية، وصيانة هذه الثروة من الضياع.

مرت المعاجم العربية بأطوار مختلفة وتعددت مدارسها المعجمية واللغوية، ومن هذه المدارس : مدرسة الخليل ومدرسة أبي عبيد ومدرسة الجوهري. ومن المعاجم المعاصرة التي سارت على نهج الترتيب الألفبائي" المعجم الوسيط" الذي أصدره مجمع اللغة العربية بالقاهرة.

إما المعجم المحوسب فهو معجم خاص يجب أن يحوي رموزًا خاصة لتصريف الكلمة ومعلومات أخرى عنها تدرج فيه مفردات اللغة بالتفصيل بحيث يكون بالإمكان الإفادة منها حاسوبيًا

أهمية المعجم العربي المحوسب

المعجم المحوسب ذو أهمية بالغة في كافة تطبيقات اللغة حاسوبيًا. فالمعجم أساس للتشكيل الآلي وللترجمة الآلية وللترجمة الفورية آليًا ولتوليد الكلام حاسوبيًا وللإملاء الآلي ولبرامج فهم الكلام آليًا ولتعليم النطق للأطفال وذوي الإحتياجات الخاصة ولتحليل النصوص ولفهم دلالة النص ولتلخيص النصوص آليًا وللإستفسار والإجابة آليًا ولكافة تطبيقات توليد الأصوات آليًا وغير ذلك من التطبيقات التي بدأت بالإنتشار بلغات أخرى وما سيظهر في المستقبل منها كثير وكثير جدًا خاصة بعد التوسع في التخاطب مع الآلات بالصوت في مختلف التطبيقات وسيكون لبصمة الصوت مستقبل يشبه بصمة الإصبع والتوقيع.

يقصد بمواصفات المعجم تحديد خصائص الكلمة العربية التي هي وحدة بناء الجملة وينبغي أن توضع لها رموز في المعجم وذلك لكي يمكن للمعجم المساعدة في كل التطبيقات التي ذكرت سابقًا أو المحتمل أن تبرز لها الحاجة مستقبلاً. وينبغي وصف الأسماء من ناحية البناء والإعراب والتذكير والتأنيث وحالة آخره من ناحية المقصور والمنقوص والممدود والتعريف والتنكير ومن ناحية العدد مفردًا أو مثنى أو جمع وإن كان جمعًا هل هو للمذكر السالم أم المؤنث السالم أم للتكسير وبالنسبة للمجرد أو المزيد هل الجذر ثلاثيًا أم رباعيًا أم خماسيًا وبالنسبة للزيادة هل هي بحرف أم حرفين أم ثلاثة وغير ذلك.

وأمّا الفعل فيأخذ الصّفات الآتية: زمن الفعل ماض أم مضارع أم أمر ومن ناحية الصحة والإعتلال هل هو صحيح سالم أم مهموز الأول أم مهموز الآخر أم معتل ناقص أم غير ذلك ومن ناحية التمام والنقصان هل هو تام أم ناقص ومن ناحية الجمود والتصرف واللزوم والتعدي هل هو متعد بنفسه أم بحرف جر وهل هو مجرد أم مزيد وعدد حروف الزيادة وهل هو مبنٍ للمعلوم أم للمجهول إلى غير ذلك من الصفات التي تلازم الفعل.

وأما الحرف فيؤخذ نوعه إن كان حرف جر أو حرف عطف أو غير ذلك وتفاصيل إعمال الحروف وهل هي متصلة أم منفصلة كتابة إلى غير ذلك من الصفات.

<u>المعاجم المحوسبة في اللغات الأخرى</u>

وبصورة عامة ليس هناك طريقة عالمية موحدة لتركيب المعجم الآلي اليوم . وقد قام بعض الباحثين بمقارنة ٥٥ معجم آلي (١٧)

تتباين المعاجم في محتوياتها تباينًا كبيرًا. وليس هناك نمط عام واحد تلتزم به المعاجم المحوسبة. فهي تختلف بين لغة وأخرى بل بين الجهات العلمية والتجارية التي قامت بوضع مواصفاتها وتغذيتها بالمعلومات اللغوية.

فهناك المعاجم أحادية اللغة التي تصف مفردات اللغة وتصريفاتها ومعانيها. وهناك المعاجم ثنائية اللغة أو متعددة اللغات. وهذه المعاجم تستند في إعدادها على المعاجم بكل واحدة من اللغات المكونة لها. لذلك فأي معجم يحوي اللغة العربية مع لغة أو لغات أخرى يحتاج إلى المعجم العربي المحوسب أحادي اللغة اساسًا.

لقد تأخر ظهور معجم عربي محوسب الآن وإن المزيد من التأخير سيدفع الشركات الأجنبية إلى أن تبني معاجم قاصرة خاصة بها نتيجة حاجاتها التجارية له. ومما لا شك فيه أن الدوافع التجارية لوحدها لن تخدم اللغة العربية كما يجب أن تُخدم. وقد بدأت بعض الشركات فعلاً بجمع الكلام العامي ليكون أساسًا لأجهزة التعرف على الكلام المنطوق مما يقلل من إستعمال اللغة الفصحى ويشيع اللهجات العامية وكتابتها مما يؤدي إلى ازدهار هذه اللهجات وتطويرها لتكون لغات مختلفة بحيث لا يفهم المتكلم بإحداها الأخرى.

ومن ناحية أخرى فإن خدمة اللغة العربية من قبل الشركات الأجنبية ينتج عنه كثير من الأخطاء اللغوية الجسيمة التي يمكن أن يقع بها أفراد من العرب ممن يقدم الخدمات لمثل هذه الشركات والمؤسسات الأجنبية وما الخلط بين ترجمة الأربعين النووية والأسلحة النووية عنا ببعيد.

محاولات الترجمة الآلية من اللغة العربية وإليها

قامت محاولات عديدة للبدء بترجمة آلية من اللغة العربية وإليها . وقد أثمر بعضها بتكوين أنظمة ترجمة آلية بينما أصبح البعض الآخر طي النسيان. وفيما يأتي بعض من هذه المحاولات:

١- محاولة الدكتور بشاي الأستاذ السابق بجامعة هارفارد ، وذلك منذ أوائل السبعينيات .

٢- برنامج ترجمان التونسي والبرامج الأخرى التي تعمل عليها عدة جهات في مصر والأردن .

٣- نظام "المترجم العربي" الذي طورته شركة ATA في لندن ، وقد طورت الشركة المذكورة برنامجا مصغرا أسمته "الوافي "

٤- نظام "عربترانز" ، وقد طورته شركة عربية أيضا في لندن.

٥- نظام "الناقل العربي" الذي طورته شركة سيموس العربية في باريس ، و لدى الشركة المذكورة أربعة برامج للترجمة بين الإنجليزية والعربية وبين الفرنسية والعربية - برنامج لكل اتجاه .

٦- نظام شركة أبتك Apptek

٧- نظام سيستران Systran برنامجا للترجمة من الإنجليزية إلى اللغة العربية وهناك موقع www.systranet.com وتقوم عليه واحدة من أشهر وأقوى الشركات في هذا المجال وتم تأسيسها عام ١٩٦٨م

٨- نظام وايدنرWeidner الذي طور أيضا برنامج للترجمة من الإنجليزية إلى العربية

٩- شركة ألبس Alps التي لا زال لديها برامج للترجمة بين عدد من اللغات ، وتطبق مبدأ الترجمة التحاورية

١٠- في فرنسا لدى جامعة غرينوبل Grenoble

١١- موقع المسبار وهو موقع يهتم أيضا بالترجمة الالية من الإنجليزية إلى العربية وبالعكس ويمتاز بالسهولة والمرونة عند استخدامه ويحقق مستوى مقبولاً من الترجمة في مستوياتها البسيطة التي أشرنا لها عند الحديث عن أنماط الترجمة .

١٢- - موقع: www.freetranslation.com

١٣- موقع: www.itranslatoronline.com .

١٤- لقد وظفت شركة صخر لبرامج الحاسب محركًا الترجمة الآلية الخاص بها في دعم موقع http://www.tarjim.com/، وهي خدمة ترجمة فورية تقوم بترجمة أي صفحة . وقد كانت أولاً ترجمة مجانية بينما الآن يمكن إرسال النص المطلوب ترجمته والحصول على الترجمة بأجور.

١٥-برنامج شركة جوجل: وهو برنامج مجاني أعلن عنه مؤخرًا ويستند إلى الترجمة الإحصائية من ذخيرة لغوية مأخوذة من الأنترنيت. والجدير بالذكر أن هذا البرنامج يتعلم من أخطائه فإذا ما ترجم جملة خاطئة وأخبره المستخدم أن الترجمة خاطئة وأن المفروض أن تكون بشكل آخر فإنه يخزن هذه المعلومات ويستعملها في المستقبل بشكل أصح ولا تزال الذخيرة اللغوية التي يستعملها محدودة لكن

مستقبل هذا النظام يتوقع له التقدم حيث أن شركة جوجل من أكبر شركات البحث على الأنترنيت اليوم.

ولتبيان القصور الذي تعاني منه المترجمات الآلية المتوفرة فقد أعطيت الجملة التالية إلى مترجمي صخر (غريب) وجوجل فكانت النتائج كالآتي:

الجملة الإنكليزية	Lawyers for the four Iraqis confirmed the first suit was registered at the U.S. District Court in Seattle, Washington
مترجم غريب	المحامون للأربعة عراقيّين أكّدوا أنّ القضيّة الأولى سُجِّلَتْ في محكمة المقاطعة الأمريكيّة في سياتل، واشنطن.
مترجم جوجل	محامون لاربعة عراقيين وأكدت الدعوى الاولى المسجله في الولايات المتحدة امام المحكمه المحلية في سياتل ، واشنطن
الترجمة الصحيحة	أكَّد المحامون عن العراقيين الأربعة أن الدعوى الأولى قد سُجِّلت في الولايات المتحدة لدى المحكمة المحلية في سياتل بواشنطن.

وهكذا يتبين أن الترجمة الآلية بوضعها الحالي لا تزال قاصرة ولكن يمكن أن تساعد المترجم الذي يعرف اللغتين أن يراجع الترجمة لوضعها بالصيغة المناسبة الصحيحة. ويبقى أن تثار استفسارات عن دلالة النص الأصلية التي قد يحار فيها المترجم من البشر هل أن النص يشير إلى أن المحامين الموكلين من قبل العراقيين قد كان تأكيدهم المذكور موجهًا للعراقيين الأربعة أم هو تأكيد صحفي عام أم غير ذلك. وربما يبرز تساؤل عن دقة ترجمة كلمة confirmed بكلمة " أكِّد" حيث أنها تعني تأييد معلومة معروفة مسبقًا بشكل غير يقيني . ويتبين أن الترجمة الآلية حتى لو وصلت بمستوى الترجمة البشرية فإنها تعاني من غموض تحتويه اللغة بكل حال من الأحوال.

جوانب أخرى ذات علاقة بالترجمة الآلية

تركز بعض الأبحاث على نطاق محدود من فوائد الترجمة الآلية: من هذه المجالات ترجمة عبارات الأستفسار بين اللغات. في هذا المجال تؤدي الترجمة الآلية لعبارة الإستفسار الى لغات أخرى ومن ثم يبحث عن موضوع الإستفسار بشتى اللغات التي تتم الترجمة اليها. ولغرض تقليل المعاني المحيرة عند الترجمة وزيادة كفاءة الترجمة يجب أن تحوي المعاجم على إشارة لموضوع الكلمة (١٨)

ترجمة الكلام المنطوق

ليس هناك ترجمة فورية آلية موثوقة لحد الآن لكن ما متوفر لا يتعدى قواميس صوتية او ترجمة لعبارات بين لغات متعددة . ويمكن ان تجد مثل هذه الأنظمة تطبيقات في حجز الفنادق والحجز على الخطوط الجوية والمشاركة في المؤتمرات وطلب الطعام من المطاعم والأستفسار عن اتجاه السير والحجز لدى العيادات الطبية والمستشفيات واستئجار السيارات وغيرها (١٩)

إن مشاكل ترجمة الكلام المنطوق تفوق مشاكل النص المكتوب نظرًا لأن هناك عادة الكثير من الكلام المتداول الذي يحوي أخطاء نحوية أو جملاً ناقصة. هذه المشاكل يمكن حل بعضها بالتحليل التركيبي المحدد. فمثلا يمكن اهمال ترتيب الكلمات في الجملة أو اهمال مفعول الكلمات المساعدة في الجملة. ويعني ذلك وضع قواعد لغوية مطاطة للترجمة الفورية الآلية.

وهذا يحدد نطاق الكلمات التي يستعملها المتكلم وطريقة نطقها وسرعة النطق وطول الجملة وصيغ التوقف بين الجمل...

هناك محاولة لاستعمال الرسائل القصيرة على الهاتف النقال للترحمة وتمتاز بأنها قصيرة ومحدودة المعجم وبأنها قابلة للتطور والتوسع(٢٠)

المشاكل والحلول

مما سبق يتبين أن هناك الكثير من المشاكل التي تعاني منها الترجمة الآلية بشكل عام وما يخص اللغة العربية بشكل خاص، ونوجزها فيما يأتي

١- عدم وجود معجم عربي محوسب
٢- قلة النصوص المترجمة بين اللغة العربية واللغات الأخرى التي يمكن الإفادة منها في تكوين ذخيرة لغوية مفيدة للترجمة الآلية التي تستند إلى أسس إحصائية
٣- قلة الأبحاث اللغوية المتعلقة بالترجمة الآلية من اللغة العربية وإليها وعدم وجود دعم كافٍ للبحث في هذا المجال ومن ذلك البحث في التحليل الإحصائي والتعرف على الأصوات ومشكلة الكلمات متعددة المعاني ومشكلة التحليل الصرفي المشترك لفظيًا ومشكلة فهم المعنى من السياق ومشكلة الإعراب والنحو ومشكلة التشكيل.
٤- عدم وجود تعاون مشترك بين الباحثين في المعالجة الآلية للغة العربية والباحثين في اللغات الشرقية الأخرى كالتركية والفارسية والأوردية والبنغالية والماليزية والسواحيلية.

إن الحلول لهذه المشاكل تتلخص بما يأتي:

١- القيام بحملة توعية للقيادات السياسية والعلمية والجهات الداعمة للبحث العلمي على أهمية البحث العلمي في حوسبة اللغة العربية على نطاق المجامع اللغوية العربية وأقسام الحاسوب واللغة العربية واللسانيات في الجامعات العربية وتوجيه الأبحاث نحو التطبيق العملي.
٢- تكوين قيادات بحثية في أقسام اللغة العربية ذات خلفية حاسوبية جيدة وفي أقسام الحاسوب ذات خلفيات جيدة باللغة العربية لكي يكون التواصل والبحث العلمي على أتم وجه.
٣- السعي لجعل اللغة العربية لغة وسيطة للترجمة الآلية بين اللغات الشرقية كالتركية والفارسية والأوردية والبنغالية والماليزية والسواحيلية والكردية ، والتواصل مع الجهات البحثية في هذه اللغات والإفادة من نتائج أبحاثها دون الإقلال من أهمية البحث العلمي في الترجمة الآلية من اللغات الأوربية إلى اللغة العربية أيضا.

الخلاصة

اللغة العربية أمانة اليوم في أعناق هذا الجيل. فالتحديات التي تواجهها الأمة في النواحي الثقافية والفكرية والتقنية كلها ذات علاقة باللغة. لذلك فإن التقاعس عن خدمة اللغة العربية اليوم يؤثر على مستقبل الأجيال القادمة وعلى نهضة الأمة وتقدمها.

الترجمة علم تطبيقي يخدم الأمة وعلى اللغويين أخذ موقف عملي للحفاظ على اللغة من جهة وتقديم الخدمة لمن يتكلم باللغة بأفضل وجه. لأن الخيار الآخر هو عزوف العامة عن اللغة واستبدالها باللغات الأخرى أو اللهجات المحلية. وعلى اللغويين إعمال جهدهم في البحث التطبيقي للغة بما ييسر تعامل اللغة العربية مع الترجمة الآلية وقد يحتاج ذلك إلى سلوك مسالك جديدة في أبحاث اللغة العربية، فإن التضحية ببعض جوانب اللغة أهون من التضحية باللغة كلها.

حين سئل الزعيم الصيني ماوتسي تونج عن رأيه بالثورة الفرنسية بعد مضي ٢٠٠ عام عليها قال : من المبكر الحكم عليها الآن. ولذلك فمن المبكر الحكم على الترجمة الآلية اليوم بعد مضي حوالي ٥٠ سنة على بدء البحث العلمي فيها.(٢١)

عند بدء البحث في الترجمة الآلية كان الهدف هو الوصول الى ترجمة آلية لكل أنواع ومحتويات النصوص وبمستوى نوعية جيدة تكافئ المترجم البشري. لكن بعد وقت قصير تبين بأن هذا الهدف لا يمكن تحقيقه في المدى المنظور لكن كذلك تبين أن كثيرًا من النصوص المترجمة ترجمة بدقة ضعيفة مفيدة للمترجمين لكي يقوموا بتحسينات عليها كما أنها تعطي فكرة عامة عن الموضوع حينما لا تكون الدقة العالية ضرورية. وحينما يصبح بالإمكان استعمال الترجمة الآلية على الأنترنيت والراديو والتلفزيون والهاتف ستكون الترجمة الآلية جزءًا من هذه المعدات وسيكون استعمالها من كل البشر. (٢٢)

إن الترجمة الآلية اليوم هي بيد الحاسوبيين واللسانيين واللغويين. لكن بعد عقدين أو ثلاثة من الزمن وفي ضوء التطور الحاصل في قدرات الحواسيب يتوقع أن تتوفر حواسيب ذات قدرات تفوق حواسيب اليوم بمئات أو آلاف المرات وعندها يكون بإمكانها القيام بعمليات تسهل عملية الترجمة الآلية كثيرًا. ولكن في ضوء العولمة والحاجة إلى التواصل بين البشر بلغاتهم المختلفة أليس من الأجدى استغلال الإمكانات المتوفرة اليوم؟ وهل استغلت جهود المترجمين من البشر كما يجب لخدمة الترجمة الآلية؟(٢٣)

نظرًا لعدم الوصول الى طريقة مثلى للترجمة لحد الآن وقد ينتظر عقدين أو أكثر من الزمن للوصول الى طريقة عامة دقيقة للترجمة الآلية فإن اعتماد أكثر من طريقة واحدة بالتوازي هو ما يحبذ الآن. وهناك أبحاث عديدة تستند إلى اعتماد طرائق متعددة بوقت واحد لترجمة النص نفسه ومن ثم يتم الإختيار بينها بعد ذلك. وتعتمد طريقة الترجمة الآلية على نوعية الترجمة المطلوبة من ناحية الدقة والرصانة اللغوية وعلى الكلفة المرصودة وعلى الزمن المستغرق وفيما إذا كان النظام قابلاً للتحسين في المستقبل أم لا.(٢٤)

المصادر باللغة العربية

(١) د.محمود اسماعيل صالح (الصيني) - أستاذ اللغويات التطبيقية - مدير مركز الترجمة بجامعة الملك سعود (سابقا) الحاسوب في خدمة الترجمة والتعريب

www.emro.who.int/ahsn/meetings/sep03/day2/Dr.%20Mohammed%20Saleh.doc
(٢) الأستاذ الدكتور محمد الصرايرة - قسم اللغة الإنجليزية- جامعة اليرموك - اللغة العربية والترجمة الآلية – محاضرة في مجمع اللغة العربية الأردني - الموسم الثقافي التاسع عشر ٢٠٠١م
(٣) د. نبيل علي - الفجوة الرقمية – عالم المعرفة ٣١٨ أغسطس ٢٠٠٥
(٤) أ.د. محمد زكي خضر " الطريقة الإجمالية في ترجمة معاني القرآن الكريم" - المؤتمر الدولي الثاني في اللغة والترجمة - مركز أطلس العالمي للدراسات والأبحاث - عمان – الأردن - ١٤-١٥/١٢/٢٠٠٢
(٥) أ.د. محمد زكي خضر "مداد البيان" - ندوة حول مشروع قاعدة بيانات حاسوبيّة للقرآن الكريم مجمع اللغة العربية الأردني ٢٠ ذي الحجة ١٤٢٧هـ ٩ كانون الثاني ٢٠٠٧م
(٦) الترجمة الآلية ... تحديات وآمال ! مجلة الحرس الوطني
haras.naseej.com/Detail.asp?InNewsItemID=62556 - 33k

المصادر الأجنبية

(7) John Hutchins, The state of machine translation in Europe and future prospects. aymara.org/biblio/mtranslation.pdf
(8) Setsuo Yamada Syuuji Kodama Taeko Matsuoka, Hiroshi Araki Yoshiaki Murakami Osamu Takano Yoshiyuki Sakamoto A Report on the Machine Translation Market in Japan. www.mt-archive.info/MTS-2005-Yamada.pdf
(9) Thepchai Supnithi, Virach Sornlertlamvanich, Thatsanee Charoenporn, A Cross System Machine Translation COLING-02 on Machine Translation in Asia – Vol 16, Sep. 2002
(10) Moustafa Elshafei1, Husni Al-Muhtaseb2, and Mansour Alghamdi3, Machine Generation of Arabic Diacritical Marks, The 2006 International Conference on Machine Learning; Models, Technologies & Applications (MLMTA'06).
(11) Nizar Habash, Bonnie Dorr, Christof Monz Challenges in Building an Arabic-English GHMT System with SMT Components, Center for Computational Learning Systems, Columbia University. www.nizarhabash.com/publications.html -

(12) Fatiha Sadat, Nizar Habash, Combination of Arabic Preprocessing Schemes for Statistical Machine Translation, Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL, pages 1–8, Sydney, July 2006.

(13) Ahmed Abdelali, Jim Cowie and Hamdy S. Soliman, Building A Modern Standard Arabic Corpus, Workshop on computational Modeling of Lexical acquisition. The split meeting Croatia 25th – 28th of July 2005.

(14) Top Ten Internet Languages – Internet Statistics.
http://www.internetworldstats.com/stats7.htm
(15) Chris Callison-Burch Philipp Koehn Miles Osborne, Improved Statistical Machine Translation Using Paraphrases In Proceedings NAACL-2006

(16) Jason Riesa, Behrang Mohit, Kevin Knight, Daniel Marcu, Building an English-Iraqi Arabic Machine Translation System for Spoken Utterances with Limited Resources. www.isi.edu/natural-language/mt/iraqi_interspeech.pdf

(17) Ali Farghaly and Jean Senellart , Intuitive Coding of the Arabic Lexicon, Proceedings of MT Summit IX, ... Jean Senellart, Director of Research and Development, SYSTRAN Software
(18) Mohammed Aljlayl, Ophir Frieder, & David Grossman, On Arabic-English Cross-Language Information Retrieval: A Machine Translation Approach. ieeexplore.ieee.org/iel5/7847/21600/01000351.pdf
(19) Ying Zhang Survey of Current Speech Translation Research, Language Technologies Institute, Carnegie Mellon University. citeseer.ist.psu.edu/752690.htm

(20) Avinash J. Agrawal, Manoj B. Chandak Mobile Interface for Domain Specific Machine Translation Using Short Messaging Service

ieeexplore.ieee.org/iel5/4151644/4151645/04151821.pdf?tp=&isnumber=&arnumber=4151821

(21) Nicola Ueffing, Hermann Ney, Word-Level Confidence Estimation for Machine Translation, RWTH Aachen University, portal.acm.org/citation.cfm?id=1220671
(22) Gerd Willée, Bernhard Schröder, Hans-Christian Schmitz (eds.) John Hutchins Computerlinguistik: was geht,was kommt? Computational linguistics: achievements and perspectives. Festschrift für Winfried Lenders (Sankt Augustin: Gardez! Verlag, 2002), p. 159-162] Machine translation today and tomorrow
www.hutchinsweb.me.uk/Lenders-2002.pdf

(23) Zhuang Xinglai, The Emerging Role of Translation Experts,in the Coming MT Era , Translation Journal, Volume 6, No. 4 Oct. 2002

(24) Fuji Ren and Hongchi Shi, Parallel Machine Translation: Principles and Practice. doi.ieeecomputersociety.org/10.1109/ICECCS.2001.930184