

الترجمة الآلية في خدمة اللغة العربية

أ.د. محمد زكي خضر

١ - مقدمة

لم تعد المعالجة الحاسوبية للغة العربية من أمور الترف اليوم، فقد أصبح واضحاً أن استعمال الآلة اليوم باللغة التي يتقنها الفرد العادي أمراً ضرورياً. وفي الوقت نفسه أصبح الحصول على كلام منطوق من الآلة أمراً متيسراً تقنياً. لذلك فإن فهم الآلة للغة العربية ومعالجتها كتابة ونطقاً هو من المهام التي ينبغي خدمتها إلى أقصى ما يمكن لكي يكون تواصلها مع البشر سهلاً ودقيقاً. إن الجهود في مجال معالجة اللغة العربية لا تزال متأخرة وهي بحاجة إلى الكثير من الأبحاث والتطبيقات. فالجهود في مجال اللغة من ناحيتي الصرف والنحو تحتاج المزيد. والموارد التي تحتاجها عملية حوسبة اللغة، هي الأخرى بحاجة إلى المعاجم الآلية والمدونات الموسومة وشبكات الكلمات والمكانز وأبحاث الدلالة. وعلى ذلك فالأبحاث في كل هذه المجالات وغيرها ضرورية لخدمة التطبيقات المختلفة التي يحتاجها التعامل اليومي مع الآلة باستخدام اللغة العربية. وأحد هذه التطبيقات المهمة هو الترجمة الآلية.

تشهد الترجمة نمواً سريعاً جداً هذه الأيام، فهي وسيلة فعّالة لنقل المعرفة والتواصل بين البشر الذين يتكلمون لغات مختلفة. وقد زاد التقدم في التقنيات الحاسوبية الحديثة وفي حقل الإتصالات هذه الأهمية وسهّل الاتصال مع الآخر. فالجانب الرئيسي للترجمة يعتمد على فهم معاني الكلمات في مفردات اللغة المراد الترجمة منها (لغة المصدر) ونقل المعنى إلى اللغة المراد الترجمة إليها (لغة الهدف). إن عملية النقل بين اللغتين يمكن أن تكون حرفية أي باستعمال معاني كلمات اللغة المصدر كلمة كلمة، وهذه الطريقة البدائية ليست هي الأفضل لأنها تفقد المفهوم العام المراد نقله من لغة المصدر إلى لغة الهدف. أما الطريقة الثانية فتهتم بالمعنى العام للعبارات في اللغة المصدر والتعبير عنها باللغة الهدف بلغة رصينة دون التقيد بحرفية النص الأصلي. أن أبرز التحدّيات التي تواجه الترجمة هو التنوع الكبير في التراكيب أو القواعد اللغوية وعملية تفسير الجملة التي تستخدم تركيباً لغوياً معيناً، ومن ثم اختيار أفضل تركيب للجملة في اللغة المستهدفة، لذلك فإن فهم المعنى العام والتعبير عنه بلغة الهدف هو التحدي الأكبر في هذه الطريقة.

تُعتبر الترجمة الآلية إحدى أهم الحُقُول وأصعبها من حيث إمكانية التطبيق في حياتنا اليومية، وهناك العديد من البحوث حول التّرجمة الآلية و تطويرها بين اللغة الإنجليزية واللغات الأخرى كالألمانية والفرنسية والإيطالية واليابانية والصينية والكورية وغيرها من اللغات. وقد أدت تلك البحوث إلى أنظمة ترجمة آلية متقدمة. لكن بالمقابل، هناك القليل من البحوث المماثلة حول اللغة العربية، لذلك فإن ما يتوفر من أنظمة للترجمة من اللغة العربية وإليها لا تزال بعيدة عن الدقة، وهي بذلك يصعب الاعتماد عليها في الوقت الحاضر.

لا بد وأن يشار إلى أن الترجمة الآلية تطورت عبر العقود القليلة الأخيرة تطورًا كبيرًا بتطور علوم الحاسوب وبتراكم الخبرات في معالجة اللغات الطبيعية التي تستخدم الذكاء الاصطناعي. كما أن تطبيق وسائل المعالجات الحاسوبية على لغة ما مكن من الغوص في دقائق اللغات المختلفة لوجود السمات العامة للغات البشرية، لكن لكل لغة خصوصياتها وهذا يحتاج إلى أبحاث خاصة لكل لغة على حدة.

إن تطوير برامج للترجمة الآلية خاصة من الإنجليزية إلى العربية أمر في غاية الأهمية لنقل ما ينشر باللغة الإنجليزية إلى العربية لكي يطلع القارئ العربي على المستجدات العلمية أولاً بأول، خاصة أن اللغة العربية ليست هي اللغة الأولى للعلوم والتقدم العلمي في العالم اليوم.

٢- تاريخ الترجمة الآلية (٢٠١)

استُخدم الحاسوب في الترجمة فعليًا لأول مرة من قبل وارن ويفر عام ١٩٤٧، فقد قدم في عام ١٩٤٩ اقتراحًا لزملائه لتطوير أنظمة الترجمة الآلية بحيث لا تقتصر على الترجمة كلمة كلمة وذلك من خلال فهم المعنى المباشر للكلمات التي تشير لأكثر من معنى واحد. وقد دقق النظر في التشابه بين تركيب المخ البشري وكيفية عمل الآلات منطقيًا. وقد توصل في نهاية بحثه إلى أن الترجمة الآلية ممكنة التحقيق. كانت هناك في تلك الحقبة حاجة سريعة للترجمة غير الدقيقة، وذلك نظرًا لحاجة الولايات المتحدة للاطلاع على المراسلات التي يقوم بها الاتحاد السوفياتي، في الوقت نفسه الذي ظهرت فيها الحواسيب وتبينت قابليتها المتنامية. ولذلك كانت أنظمة الترجمة من الروسية إلى الإنجليزية أولى المحاولات، فقد أجريت بنجاح في جورج تاون أول ترجمة من اللغة الروسية إلى الإنجليزية عام ١٩٥٤، وقد تنبه الباحثون في الإتحاد السوفياتي إلى أهمية الترجمة الآلية فقاموا في عام ١٩٥٥ بأول تجربة في الترجمة الآلية من الإنجليزية إلى الروسية باستعمال قاموس يحوي ٢٣٠٠ كلمة. وأشارت التجربة إلى ضرورة تمثيل التراكيب اللغوية بمستوى بناء الجملة وبالمستوى المعجمي. وقد اتضح أن مشكلة الغموض في التعامل مع الترجمة الآلية هي

مشكلة فعلية لكنها لم تقدر حق قدرها حينئذ. لقد كانت الفترة من ١٩٥٦ إلى ١٩٦٦ فترة التوقعات الكبيرة من الترجمة الآلية لكن لم يتحقق ما كان يؤمل من الحصول على دقة بنسبة ٩٥% دون معالجة مسبقة للنصوص، فلقد أوضحت الأبحاث أن تركيب اللغات أكثر تعقيداً مما كان متصوراً، وأن المترجمين من البشر يستعملون الكثير من المعرفة غير المنظورة في حقل المادة المترجمة والمعرفة بالحياة العادية أكثر مما كان متصوراً. وأدى ذلك إلى تباطؤ في أبحاث الترجمة الآلية وربما إهمال الموضوع حتى عام ١٩٧٥.

لقد كانت أولى محاولات تطوير برنامج ترجمة من الإنكليزية إلى العربية في نهاية السبعينيات في ولاية يوتا الأمريكية، ولم يكن يحتوي البرنامج على تحرير أولي، لكن كان بالإمكان إجراء تحرير نهائي. وقد احتوى النظام على مرحلتين الأولى لتحليل اللغة المصدر والثانية لتوليد اللغة الهدف. كان تحليل اللغة الإنكليزية موجهاً لكي يولد العبارات في اللغة العربية بشكل مباشر باستعمال معجم ثنائي اللغة للعبارات. ولم تكن الألفاظ والجمل الإنكليزية تحلل بعمق، بل كان يكفي توليد المقابل باللغة العربية. وهكذا كان النظام باتجاه واحد ولم يكن يغوص بعمق في اللغة الإنكليزية. وقد استعمل البرنامج نفسه فيما بعد لترجمة الموسوعة البريطانية (بريتانكا) إلى اللغة العربية. كما قامت سلطنة عمان بالحصول على امتياز للبرنامج لكي يستعمل في ترجمة الوثائق الإنكليزية إلى العربية. لم تكن الطريقة المباشرة هذه في الترجمة تغوص في أعماق اللغة المصدر، ومن ثم أدرك مطورو البرمجيات أن الطريقة المباشرة لا تكفي لمعالجة اللغات الطبيعية المعقدة. ولنضرب مثلاً على إحدى الصعوبات التي جوبهت: فقد وجد أن التعامل مع الجمل باستبدال مواقع الفعل بين الإنكليزية والعربية مثل "Ali drinks tea" فترجمتها إلى: "يشرب علي الشاي" احتوى على استبدال موقع الفعل والفاعل بحيث أصبحت الترجمة العربية جملة فعلية حيث تتبدىء بالفعل بدل الاسم. لكن ذلك لم يكن كافياً لترجمة جملة مثل :

The man whose clothes are dark and carrying a suitcase in his left hand ran away.

فمثل هذه الجملة تحتاج إلى تحليل عميق لتحديد الفاعل ومن ثم اقترانه بالفعل الذي يبعد عن الفاعل بأكثر من عشر كلمات.

ومن المشاكل التي تعاني منها الترجمة مدى التوافق في حل مشكلة ترتيب العبارة التي تحوي على أداة النفي not والتوافق بين الصفة والإسم والمحتوية على did not فيما يتعلق بالجنس والحالة، والتنكير والتعريف، والضمائر الملحقة، والتوافق بين الفعل والفاعل، وعملية الإضافة والحذف بين اللغتين. يظهر ذلك إذا كانت اللغة الأصلية تحتوي على كلمات إضافية خاصة بها ليس لها وجود في اللغة الهدف مثلاً.

وعلى ذلك فقد أصبح واضحًا أن الطريقة المباشرة لا تعطي نتائج دقيقة لترجمة مثل هذه الجمل الطويلة. وهكذا برزت حاجة لتطوير المعرفة بكيفية تحليل الجمل الطويلة بعمق وكفاءة دون غموض، وعند ذلك ظهرت الطريقة التحويلية التي قدمت للترجمة الآلية أمرين: الوصف لتحليل الجملة ومن ثم التقنية الجديدة لتمثيل هذا التحليل العميق، وهو ما طور خلال السبعينيات والثمانينيات من القرن الماضي.

عادت الأبحاث على أشدها في أوروبا وكندا واليابان بين عامي ١٩٧٥ و ١٩٨٥ باستعمال الأنظمة الخبيرة وأبحاث معالجة اللغات الطبيعية، وبشكل خاص للغات الأوربية واليابانية، وبذلك ظهر الجيل الثاني من برامج الترجمة الآلية فظهرت في الأسواق برامج تجارية تستعمل الحواسيب المايكروية. وهذا التطور شمل البحوث المعجمية والنحو والصرف والدلالة.

وفي عام ١٩٨٩ كانت بداية الجيل الثالث من برامج الترجمة الآلية التي استندت إلى المعلومات الإحصائية حينما قامت شركة أ.ب.م.، بمشروع كارديد المستند إلى الترجمة بالأمثلة والترجمة المحدودة الموضوع مع تعدد اللغات المترجم منها والمترجم إليها، فعاد الاهتمام بالترجمة الآلية لحد لم يسبق له مثيل. وقد استندت هذه الترجمة إلى الذخيرة اللغوية Corpus Based MT والتي لا تزال في تطور حتى اليوم مع بعض التكامل مع الأساليب الأخرى للترجمة الآلية.

لقد حدث تقدم هائل في برامج الترجمة الآلية بين اللغات الأوربية فيما بينها وبين اليابانية والصينية والكورية، وقد ازداد الطلب عليها بشكل واسع وساهمت الشبكة بزيادة الحاجة إلى الترجمة الآلية وسهّلت في تبادل المادة المترجمة إلى من يحتاجها. ولا يزال التقدم على أشده في تكامل عمل المترجمين من البشر مع الترجمة الآلية، حيث يحتاج التقدم بمستوى الترجمة الآلية إلى مترجمين أكفاء لتطويرها والارتقاء بها، وليس كما كان يخشى بأن الترجمة الآلية ستؤدي إلى انقراض مهنة الترجمة البشرية .

إن استعمال أنظمة الترجمة الآلية قليلة الدقة يدفع إلى الحاجة إلى تحسين النوعية بشكل مطّرد بمساعدة المترجمين من البشر. وقد وظفت بلدان كثيرة الترجمة الآلية لخدمة متطلباتها الاجتماعية والاقتصادية والعلمية والتقنية في طليعتها الولايات المتحدة وروسيا والصين واليابان. وفي الحقل العلمي هناك جهود كبيرة لترجمة المحاضرات العلمية المتوفرة مجانًا على الشبكة إلى لغات أخرى. إن أكبر مجموعة محاضرات متقدمة مفتوحة المصدر متوفرة على الشبكة هي محاضرات جامعة إم آي تي الأمريكية. وهي الآن تحظى بترجمة آلية إلى اللغة الصينية مع مراجعة بشرية لما بعد الترجمة الآلية. (٣)

تجري منذ الثمانينيات في القرن الماضي العديد من الأبحاث حول المعالجة الصرفية للغة العربية وكان معظم تلك الدراسات ينصب على عملية فصل السوابق والواحق في الكلمة العربية المركبة،

ومن ثم استخلاص الجذر لمعرفة العمليات الصرفية التي أجريت عليها والتي يمكن للحاسوب القيام بها بعد برمجته بشكل دقيق. وتجدر الإشارة إلى أن قواعد الصرف العربية مهياة للبرمجة أكثر من قواعد النحو التي تحتاج إلى ارتباط بالمعنى والدلالة أكثر من الصرف.

٣- أنواع الترجمة الآلية

٣-١- الترجمة الآلية المستندة إلى المعاجم والقواعد

من البديهي أن تكون المحاولات الأولى للترجمة الآلية مستندة إلى المعاجم ثنائية اللغة ثم بعد ذلك المضي بإدخال قواعد صرفية ونحوية ودلالية على الترجمة. وقد وجد أنه من الصعوبة بمكان فصل العمليات الصرفية والنحوية والدلالية عن بعضها البعض. كما أن المزج بينها بطريقة علمية أمر في غاية الصعوبة. ومع هذا تحقق بعض التقدم في محاولات ترجمة الجمل القصيرة ذات التراكيب البسيطة الشائعة.

٣-٢- الترجمة الآلية الإحصائية

"الترجمة الآلية الإحصائية" هو مصطلح يشير إلى مجموعة من أنظمة الترجمة الآلية التي طورت باستخدام أساليب التعلّم الآلي، وهي النوع الأكثر شيوعاً الآن من أنواع الترجمة الآلية. لقد طُورت أولى أنظمة الترجمة الآلية الإحصائية قبل ما يقرب من ثلاثة عقود، واستخدمت شركة IBM في بداية ظهور الترجمة الإحصائية نظرية Bayes لتطوير منهجية إحصائية للترجمة الآلية، حيث اعتمدت على سلسلة من الكلمات والرموز الموجودة في لغة المصدر مع مجموعة من المفردات يمكن أن تُسمى (أ) وتحويلها إلى سلسلة من الكلمات والرموز الموجودة في اللغة الهدف مع مجموعة أخرى من المفردات تسمى (ب).

تتميز أنظمة الترجمة الآلية الإحصائية بقدرتها على تعلّم ترجمة العبارات وليس الكلمات المفردة فقط، وميزة أخرى لبعض هذه الأنظمة أنها تجمع عدة موارد معرفية لاستخدامها لتحليل الجملة المترجمة. ويجري وصف الترجمة الآلية الإحصائية عادة بأنها اختيار الجملة (أو العبارة) الهدف التي لها أعلى احتمالية مقابل الجملة (أو العبارة) المصدر.

قام كيفن نايت ودانيال ماركو بتحويل الأبحاث العلمية المتعلقة بالترجمة الآلية إحصائياً في عام ٢٠٠٢ إلى منتج للترجمة الآلية من العربية إلى الإنكليزية، وكان ذلك مثلاً جيداً لتكوين ذخيرة

لغوية متقابلة بين العربية والإنكليزية والاستفادة منها للترجمة الآلية. وأشهر الأمثلة على الترجمة الإحصائية بين اللغة العربية واللغات الأخرى هي ترجمة جوجل.

٣-٣- الترجمة الآلية باستخدام الشبّكة الدلاليّة

تعمل الشبّكة الدلالية على إعادة هيكلة كمّ هائل من البيانات التي يمكن الوصول إليها وتكون مفهومة لكل من البشر والآلات المتاحة على الشبّكة بطريقة مشابهة لتلك التي يُدركها العقل البشري، وتكون بمثابة تدريب للشبكة على فهم السياق القريب من أي كلمة أو عبارة يتمّ البحث عنها.

٣-٤- الترجمة الآلية العصبية

تُعتبر الترجمة الآلية العصبية موضوعاً جديداً، بدأ يشهد في الآونة الأخيرة نشاطاً، حيثُ تجري خلالها الترجمة الآلية بشكل مختلف تماماً عن الطرق التقليدية لأساليب الترجمة الآلية الإحصائية القائمة على العبارة. فبدلاً من التمرّن على المكونات المختلفة للترجمة الآلية بشكل مستقل، يستخدم هذا النموذج الشبكة العصبية الاصطناعية لتعليم النموذج أو تدريبه على جمع المكونات معا لتحقيق أكبر قدر من الإتقان في أداء الترجمة بواسطة خطوتي الشبكة العصبية المتعاقبتين: "التشفير" و "فكّ التشفير".

٣-٥- الترجمة الآلية الهجينية

تمكّنت الترجمة الآلية من حلّ الإشكال في بعض الأحوال باستخدام معلومات إضافية مثل الترجمة الآلية المعتمدة على القواعد التي تساهم في تصحيح الأخطاء باستخدام معلومات دلالية، بينما تستخدم أنظمة الترجمة الآلية الإحصائية الإحصاء لكمية هائلة من الذخيرة اللغوية ثنائية اللغة أو متعددة اللغات. وكلا النظامين له نقاط قوة وضعف، وقد ساهما بشكل منفصل بحل جزء من المشكلة، أما الترجمة الآلية الهجينة فزادت من تحسين المخرجات لأنظمة تمييز الكلام الآلية (ASR) (automatic speech recognition) وأنظمة التحكم بالوسائط بالجمع بين أهم ميزات النظامين. وقد أثبتت التجربة أن النظام الهجين يكون أدائه أفضل من النظام الإحصائي وحده أو الطرق المعتمدة على القواعد وحدها.

كما ظهرت بعض الأنظمة الهجينية التي تستعمل للترجمة المكتوبة والمنطوقة باستعمال الطريقتين: الطريقة الإحصائية والطريقة المستندة إلى القواعد، حيث أن لكل من الطريقتين فوائدها ونقائصها. لذلك فمن المزاجية بينهما يمكن الحصول على الفوائد من كليهما واستبعاد النقائص، وهو أمر ليس بالهين من ناحية التطبيق خاصة في الترجمة الآلية الفورية (للكلام المنطوق). ولغرض المضي في تحسين أداء الترجمة الآلية من العربية وإليها يجب القيام بعمليات مختلفة منها اكتشاف أسماء الأعلام ببرنامج خاص وهذا يمكن أن يحسن من الترجمة، فما أن يكتشف مثل هذا الاسم حتى يمكن القيام بعدة عمليات تفيد في الترجمة، فمثلا عند اكتشاف أن كلمة مثل "أحمد" هي اسم علم فإنها لا تترجم على أنها فعل مضارع يعود إلى المتكلم (أحمدُ I praise).

يوظف نموذج الترجمة الآلية المعتمد على القواعد نظاما يتوفر فيه معجم غني بالشروح يحتوي بدوره على معلومات دلالية ووظيفية، ويُستخدم النظام لتغذية عمل الترجمة الآلية الهجينة.

٤- الترجمة الآلية من العربية وإليها:

إن أحد المشاكل الرئيسية في الترجمة من الإنكليزية إلى العربية هو : التوافق وإعادة الترتيب بين الصفات والأسماء والتوافق بين الفعل وفاعله والتوافق بين الضمائر وأسماء الموصول، إذ أن ترتيب الكلمات في هذه المجالات الثلاثة مختلف بين العربية والإنكليزية.

فتحويل موقع الفعل والفاعل والمفعول مثل: "The patient needs a special treatment" إذ تصبح "احتاج المريض معالجة خاصة" وذلك بتقديم الفعل على الفاعل في العربية. وتحويل أداة التعريف the لما قبل الصفة مثل "The effective health treatment" وتكرار ألف لام التعريف وإعادة ترتيب الصفة والموصوف فتصبح "المعالجة الصحية الفعالة".

يرى البعض بأن أبرز التّحدّيات التي تواجه التّرجمة هو التّنوع الكبير في التراكيب أو القواعد اللّغويّة وعملية تفسير الجُملة التي تستخدم تركيباً لغوياً معيّنًا، ومن ثمّ اختيار أفضل تركيب للجُملة في اللّغة المُستهدفة.

لقد أجريت أبحاث عديدة لدراسة ترتيب الجمل المحتوية على الفعل بين الإنكليزية والعربية. وبعد ذلك استعمال هذا الترتيب في الترجمة الآلية الهجينية لتحسين نوعية الترجمة. تحاول الترجمة المستندة إلى الأسس التحويلية إيجاد تمثيل وسيط يلتقط المعنى الأصلي لكي يحوله إلى اللغة الهدف، فالغرض هو اكتشاف كيفية تحويل الجمل المحتوية على (فاعل – فعل – مفعول به) من الإنكليزية

إلى (فعل – فاعل – مفعول به) بالعربية. وتحاول الترجمة المستندة إلى الأمثلة التعامل مع الحالات غير النظامية وذلك لكي تحاكي الترجمة المثل المعروض. لقد ظهرت عدة برامج ترجمة من العربية وإليها، منها: الكافي وصخر وجوجل وبينغ وسيستران، وحاولت كثير من الأبحاث مؤخرًا مقارنة عدد من برمجيات الترجمة. وقد خلص أحد هذه الأبحاث البحث إلى وجود ١٢ مشكلة تواجه برامج الترجمة الآلية وتحدد دقة الترجمة. من هذه المشاكل: مدى التوافق في حل مشكلة ترتيب العبارة التي تحوي على أداة النفي not والتوافق بين الصفة والإسم والمحتوية على did not فيما يتعلق بالجنس والحالة والتنكير والتعريف والضمائر الملحقة والتوافق بين الفعل والفاعل وعملية الإضافة والحذف بين اللغتين. ويزداد ذلك صعوبة إذا كانت اللغة الأصلية تحتوي على كلمات إضافية ليس لها وجود في اللغة الهدف مثلاً. وقد كانت بعض الأبحاث تركز على ترجمة في حقول معنية كالترجمة الآلية في حقل الزراعة وذلك لضمان دقة أفضل في الترجمة خاصة إذا كانت بحقل علمي ضيق.

٥- نماذج لأبحاث ودراسات للترجمة الآلية إلى اللغة العربية (٤١)

أجريت العديد من الأبحاث لتعزيز قابليات أنظمة الترجمة الآلية إلى اللغة العربية مثل اعتماد أنظمة الترجمة على ثلاثة مجموعات من القواعد: التحليل النحوي والتحويل والتوليد. وأجريت تجارب على ترجمة لعدد محدود من الجمل من مجلة طبية متخصصة ومن ثم الحصول على دقة لا بأس بها، كما أجريت تجارب على عملية عكس تسلسل الأسماء مع صفاتها. وأجريت تجارب على تقسيم النص إلى جمل ثم تقسيم الجمل الطويلة إلى جمل قصيرة ثم تقسيمها إلى أجزاء من حروف جر وحروف عطف وغيرها وترتيب الجملة بشكل مصفوفة مناسبة لإعادة الترتيب بين الإنكليزية والعربية، ومن ثم تقسيم أجزاء الجمل إلى عبارات ومكونات العبارات من أسماء وأفعال وصفات وظروف وبعد ذلك تكوين شجرة إعرابية للجملة الإنكليزية.

أما ما يتوفر في السوق من برامج، فقد قارنت بعض الدراسات الترجمات المختلفة المتوفرة في الأسواق كترجمة جوجل وبابلون وتبين بأن ترجمة جوجل أفضل من ترجمة بابلون. كما قورنت أنظمة ترجمة جوجل وبابلون وسيستران وفق صفات لغوية وتبين بأن جوجل هو الأفضل كما قورنت الترجمات عبر السنوات فوجد تحسن في هذه الترجمة بين ٢٠٠٦ و ٢٠١٦. وتعتبر الإنكليزية لغة وسيطة للترجمة من عدة لغات إلى العربية (كالروسية مثلاً) حيث أن الترجمة من تلك اللغات (إلى الإنكليزية) قد بلغت دقة عالية ومن ثم تحتاج إلى تحسين الترجمة من الإنكليزية إلى العربية لترتقي بالترجمة من تلك اللغات إلى العربية بشكل غير مباشر. إن تحسين الترجمة إلى العربية لا يعتمد فقط على الحاسوبيين بل يعتمد على اللسانيين الذين تقع على عاتقهم وضع اللغة العربية بقوالب وقواعد حاسوبية تمهيدا لتضمينها في برامج الترجمة الآلية.

كما قامت بعض الأبحاث بجمع أكثر من ٦٠٠٠ جملة وتوزيعها عشوائياً على ٥٠ محكماً بمعدل حوالي ٢٥٠ جملة للمحكم الواحد لقياس التباين بين المترجمين وجرى تقييم ١٠ تقديمات يدوياً بمعدل حوالي ٤٠٠ جملة، لذلك فإن حوالي ٨٠٠٠ جملة قيمت مرتين ومن ثم استخلص أن عملية التدريب أدت إلى تحسين النتائج لكن ليس بالمستوى المطلوب، حيث أن البيانات لم تكن كافية.

كما جمع باحثون آخرون مجموعة مقالات تحتوي على مائة ألف كلمة من المقالات الإخبارية الإنجليزية مأخوذة من الموقع الإلكتروني للصحافة التعاونية على الشبكة، وقد ضمت المجموعة خمسمائة وعشرين مقالة، بمعدل مائة واثنين وتسعين كلمة لكل مقالة. ترجمت تلك المجموعة تلقائياً باستخدام خدمات جوجل المدفوعة الأجر للترجمة. وللحصول على نسخة منسقة بعد تحريرها من أخطاء الترجمة الآلية من قبل مجموعة المدققين، برزت الحاجة إلى إرشادات تصحيحية واضحة وموجزة. وقد وضع الباحثون الأخطاء المراد تصحيحها في فئات الأخطاء التالية: التدقيق الإملائي، إختيار الكلمات، أصل الكلمات، النحو، الأسماء الصحيحة، استخدام اللهجات وعلامات الترقيم. وأوضح الباحثون في المبادئ التوجيهية، أنه ينبغي تصحيح النصوص المترجمة الآلية بأدنى عدد من التعديلات اللازمة لتحقيق جودة مقبولة للترجمة، وقد طلبوا من المدققين إيلاء الاهتمام إلى الجوانب الثلاثة التالية: الدقة والطلاقة والأسلوب من أجل إنتاج نص متناسق. وقد أوضحت الأقسام التالية لهذا البحث إجراءات الشرح وإجراءات مراقبة الجودة باستخدام تدابير متكررة بين الشروح. وقد خلص البحث إلى أن هذه المنهجية المتعلقة بوضع المبادئ التوجيهية والتحقق من إتساق الشروح يمكن تطبيقها في مشاريع أخرى ولغات أخرى أيضاً.

وفي بحث آخر جرى بناء ذخيرة لغوية بتحرير لاحق للترجمة الآلية العربية، استعرضت في البحث القواعد الإرشادية مع أسلوب التعليقات (annotation procedure) لتكوين ذخيرة لغوية يجري عليها تصحيح لاحق من البشر خاصة باللغة العربية القياسية الحديثة. وقد استخدم فريق عمل مكون من أشخاص جرى تدريبهم ثم قاموا بوضع هذه التعليقات. وكانت هذه الذخيرة جزءاً من بنك قطر للغة العربية الذي هو مشروع ضخم لعمل التعليقات اليدوية. وهدف هذا المشروع هو تكوين ذخيرة تحوي مليوني كلمة للمستخدمين عبر الشبكة لجمع ملاحظاتهم على مواقع الأخبار وغيرها.

وقد اختير في تلك المرحلة جزء مقداره مائة ألف كلمة بهدف الترجمة الآلية من مختلف المواقع الإخبارية الإنكليزية المترجمة للعربية باستعمال ترجمة جوجل كمرحلة أولى.

وفي بحث آخر عن تمييز الكلام المحكي باللغة العربية العامية من خلال تحويله إلى اللغة العربية المعاصرة أُنتمت ترجمته بعد ذلك إلى الإنكليزية سُجّلت المدخلات من قناة تلفزيونية. إن سماع الكلام المحكي وتحويله إلى نص مكتوب ثم ترجمته هما عمليتان منفصلتان، وعندما يكون هناك خطأ في تحويل الكلام إلى نص فستكون الترجمة خاطئة. وقد استخدم في الترجمة الآلية الاعتماد على القواعد التي تساهم في تصحيح الأخطاء باستخدام معلومات دلالية.

٦- واقع أنظمة الترجمة الآلية العربية

سنقارن بين ثلاث أنظمة للترجمة الآلية هي جوجل وسيستران وبينغ باستعمال مواقعها على الشبكة. وقد اخترنا جملة عربية لترجمتها ومقارنة ترجماتها إلى الإنكليزية. ثم ترجمنا جملة عربية إلى الإنكليزية وأدخلت تلك الترجمة للأنظمة الثلاثة للحصول على الترجمة العربية. وقد أعيدت التجربة مرتين في شهري أيلول (أشير لها بالرمز (*)) وكانون الأول (أشير لها بالرمز (**)) من عام ٢٠١٤. ثم أعيدت التجربة في تشرين الأول ٢٠١٨ (أشير لها بالرمز (***)) ويلاحظ حدوث اختلاف في كل الترجمات من الإنكليزية إلى العربية في المواقع الثلاثة، بينما الترجمة من العربية إلى الإنكليزية اختلفت في ترجمة جوجل وبقيت نفسها في سيستران وبينغ. ويلاحظ عدم الدقة والركاكة في الترجمات الآلية كلها تقريباً. كما يلاحظ أن الترجمة قد تختلف بمضي الزمن.

النص الأصلي بالعربية	حصل تقدم هائل للترجمة الآلية خلال السنوات الماضية خاصة بين اللغات الأوروبية ذات الأصول اللاتينية لما بينها من مشتركات
ترجمة جوجل إلى الإنكليزية	Got tremendous progress of the translation mechanism during the past years, especially among European languages with Latinos because of the participants, including (*)

<p>Progress has been made tremendous machine translation during the past years, especially among European languages with Latin asset to them from participants(**)</p> <p>There has been tremendous progress in machine translation over the past years, especially among European languages with Latin origins, among them participants (***)</p>	
<p>Huge progress for the automatic translation during the last years collected especially between the European languages self of the Latin origins for what between her from joint (*) & (**) & (***)</p>	<p>ترجمة سيستران إلى الإنكليزية</p>
<p>Advances of machine translation over the past years, especially between the European languages with Latin origins to their subscribers (*) & (**)</p> <p>Huge progress of automatic translation through the years, especially among European languages Latin origins to their subscribers (***)</p>	<p>ترجمة بينغ إلى الإنكليزية</p>
<p>A tremendous progress occurred during recent years between European languages due to their common Latin origin.</p>	<p>ترجمة النص بالإنكليزية</p>
<p>حدث تقدم هائل خلال السنوات الأخيرة بين اللغات الأوروبية بسبب الأصل اللاتيني المشترك(*) & (***)</p> <p>حدث هناك تقدما هائلا خلال السنوات الأخيرة بين اللغات الأوروبية بسبب الأصل اللاتيني المشترك(**)</p>	<p>ترجمة جوجل إلى العربية</p>
<p>تقدم هائل وقع أثناء سنوات الأخيرة بين لغات أوربيّ واجب إلى أصلهم لاتيني عادي (*)</p>	<p>ترجمة سيستران إلى العربية</p>

تقدم هائل وقع أثناء سنوات الأخيرة بين لغة أوربيّ واجب إلى أصلهم عاديّ لاتينيّ (***) و (***)	
التقدم هائل الذي حدث خلال السنوات الأخيرة بين اللغات الأوروبية بسبب أصلها اللاتيني الشائعة (*) & (***) التقدم هائل الذي حدث خلال السنوات الأخيرة بين اللغات الأوروبية بسبب أ صلهم اللاتينية الشائعة (**)	ترجمة بينغ إلى العربية

يلاحظ الخطأ في التمييز بين المذكر والمؤنث في الترجمة من الإنكليزية إلى العربية في ترجمتي سيستران وبينغ، بينما ترجمة جوجل عانت في مرحلة ما من الخطأ بين الرفع والنصب.

أما الترجمة من العربية إلى الإنكليزية فيلاحظ أن ترجمة سيستران لم يلاحظ عليها تغير مع الزمن إذ كانت الترجمات الثلاثة نفسها ، بينما كان هناك بعض التحسن في ترجمة بينغ مع الزمن وكذلك ترجمة جوجل التي كانت أفضل الأنظمة الثلاثة. لكن يلاحظ أن الترجمات الثلاثة تعاني من ركاكة في التعبير وعدم دقة في تبيان المقصود.

٧- متطلبات تطوير الترجمة الآلية من اللغة العربية وإليها

٧-١- المحللات اللغوية العربية (Arabic Parsers)

إنّ محلّ اللغة الطبيعية هو برنامج يقوم بعملية تشبه عملية إعراب الجملة نحوياً حيث يمكنه أن يفهم أجزاء الجملة ويساعد في الترجمة باستخدام الترجمة الآلية، وقد قام العديد من الباحثين بدراسة هذا النوع من المحلّلات. وهناك حاجة للمزيد من البحث في هذا المجال

٧-٢- المحللات الصّرفية (Morphological Analyzer)

يعتبر الغموض الصّرفي مصدر قلق بالنسبة للمحللات النّحوية، وغيرها من أدوات معالجة اللغات الطبيعية. ويعطي التّحليل الصّرفي معلومات أدق حول أجزاء الكلام بحيث يختار التّحليل الأنسب لها بشكل يتوافق مع السّياق.

٧-٣- المحلّلات الدلالية

المسألة الأساسية التي يجب أن تُعالجها أنظمة الترجمة الآلية هي فك الغموض الذي يكتنف كثيرا من التعابير في اللغة المصدر وتحويلها إلى جمل واضحة ومفهومة بلغة الهدف. تعتبر الأنطولوجيا هي الطريقة الملائمة للقيام بكل ذلك بوضع مواصفات صريحة للمفاهيم.

٧-٤ - إعادة ترتيب الجملة (خوارزميات التصحيح)

يجب أن تضمن الترجمة الآلية دقة ترجمة اللغة المستهدفة، سواء من العربية أو إليها. لذلك، في كلتا الحالتين، يجب التحقق من تركيب الجملة. فاللغة العربية تتضمن تراكيب مختلفة، مثلا بحيث يحل الاستبدال التلقائي محل الجنس اللفظي وترجمته النسبية وفقاً للسياق.

٨- المقترحات المستقبلية

إن مسألة الترجمة الآلية من اللغة العربية وإليها لم تحل لحد الآن. فهناك حاجة للكثير من البحث العلمي في هذا المجال لكسب الطمأنينة إلى الوصول إلى مستوى من الترجمة يقترب من الترجمة من قبل مترجمين من البشر. إن الحاجة إلى ترجمة أسرع وأرخص بين اللغات سنتعزز فقط بالمشاركة بين المعلومات بين الأمم. واستناداً إلى ذلك نقترح ما يأتي:

هناك حاجة لاستعمال كمية كبيرة من الذخيرة اللغوية (corpus) المتوازية (بلغتين أو أكثر) لغرض تدريب الترجمة الإحصائية للحصول على نتائج أفضل. وهو ما لا يتوفر الآن بين اللغتين الإنكليزية والعربية، حيث يجب استعمال كمية كبيرة من البيانات في كل حقل من حقول البيانات المراد الترجمة فيه. وكلما قمنا بتوليف (tuning) النموذج بشكل أفضل كلما كانت النتائج أفضل، لذلك يجب القيام بعملية التوليف بقدر الاستطاعة. كما يجب استعمال بيانات الفحص في حقل المعلومات المراد الترجمة فيه نفسه للحصول على نتائج أفضل. إن التكامل مع وسائل أخرى للترجمة الآلية مثل الشبكات العصبية الذكية (artificial neural networks) في عملية التعلم وعملية التدريب للحصول على نتائج أفضل ضروري ويجب الاستمرار في البحث فيه. كما أن من الضروري العمل على المعجم العربي الإلكتروني التشابكي بما يشبه wordnet حيث أن ذلك في غاية الضرورة وما يتوفر باللغة العربية AWordnet لا يعتبر كافياً البتة. هناك بالطبع الحاجة إلى المزيد من الأبحاث في مجالات الصرف والنحو والأنطولوجيا وتركيب الجملة العربية.

٩- الاستنتاجات والمقترحات

باتت الترجمة مؤخرًا نشاطاً تجارياً واسعاً، خاصة في ظل تسارع وتيرة النمو في قطاع تكنولوجيا المعلومات والاتصالات وهذا يتطلب تطوير الآلات والأدوات التي تُسهّل التواصل بين الأطراف المختلفة.

والترجمة الآلية هي واحدة من تلك الأدوات اللازمة؛ بل إنها اعتُبرت الجُندي المجهول في أي عملية من عمليات الاتصال. ومع ذلك، لم تُفلح الترجمة الآلية -حتى اللحظة- في التفوق على الترجمة البشرية. لكن لحسن الحظ، لم يتوقف البحث العلمي في هذا المجال، ونأمل أن تستمر مساهمته. وفي بعض الأحيان، قد تكون هناك حاجة إلى بعض الأدوات الأخرى لاستخدامها في الترجمة الآلية، للمساعدة في تحليل النص ومن ثم تطوير النصوص في اللغات المُستهدفة. ولا تزال هذه البحوث مُستمرة ولن تتوقف أبداً حتى تُصبح الترجمة الآلية مُكافئة أو أفضل من ترجمة الإنسان.

الترجمة الآلية من اللغة العربية وإليها إحدى الوسائل التي تكمن فيها خدمة كبيرة في نهضة الأمة وأجيالها المستقبلية لأن اللغة العربية اليوم ليست هي اللغة العالمية للعلوم والتقنية، فما يصدر في اللغات الأخرى وخاصة الإنكليزية من أبحاث ودراسات وكتب ومؤلفات أخرى يحتاج للترجمة للغة العربية. لذلك فإن تطوير الترجمة الآلية من الإنكليزية إلى العربية يعتبر من أولى الأولويات في نهضة الأمة العربية. كما أن اللغات التي يتكلم بها المسلمون بحاجة إلى التراث العربي الإسلامي غير المتوفر في تلك اللغات وهناك حاجة ماسة لترجمته. لذلك فالترجمة الآلية من اللغة العربية وإليها ليست من الكماليات بل هي من الضرورات القصوى لنهضة الأمة وعودتها لمكانتها بين الأمم. وأول ما تحتاجه الترجمة الآلية هو وجود معجم عربي محوسب وتكوين ذخيرة متعددة اللغات مع العربية لتساعد في الترجمة الآلية المستندة إلى الإحصاء وما يتعلق بها من ترجمة هجينية، بدعم الأبحاث اللغوية المتعلقة بالترجمة الآلية من اللغة العربية وإليها. ولغرض الوصول إلى ذلك هناك حاجة إلى القيام بحملة توعية للقيادات السياسية والعلمية والجهات الداعمة للبحث العلمي على أهمية البحث العلمي في حوسبة اللغة العربية من قبل الجامعات اللغوية العربية وأقسام الحاسوب واللغة العربية واللسانيات في الجامعات العربية ومراكز البحث العلمي، كما ينبغي توجيه الأبحاث نحو التطبيق العملي وتكوين 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. لذلك يقع على كاهل مراكز الأبحاث والجهات الداعمة لها تكوين موارد لغوية مفتوحة المصدر ومشاع استعمالها من ذخيرة لغوية متعددة اللغات ومعاجم لغوية إلكترونية وتطوير برامج الصرف

والتشكيل الآلي وتوليد الكلام وغيرها خدمة للترجمة الآلية، هذا بالإضافة لما يقع على كاهل شركات أنظمة معالجة البيانات من تطوير للبرمجيات وتحويلها إلى تطبيقات عملية.

١٠- المراجع:

١- محمد زكي خضر : الترجمة الآلية من العربية وإليها ، الحرف العربي والتقنية – أبحاث في الحوسبة العربية – مركز الملك عبدالله بن عبد العزيز الدولي لخدمة اللغة العربية ٢٠١٤

٢- Nabeel T. Alsohybe, Neama Abdulaziz Dahan and Fadl Mutaher Ba-Alwi
Machine-Translation History and Current Evolution: Survey for Arabic-English Translations
Journal of Applied Science and Technology, 23(4): 1-19, 2017;
Article no.CJAST.36124

٣- <https://ocw.mit.edu/courses/translated-courses/traditional-chinese>

٤- محمد زكي خضر، الترجمة الآلية إلى اللغة العربية ضرورة ملحة – المؤتمر الدولي السابع للغة العربية – دبي ١٩-٢٢ أبريل ٢٠١٨

5- Hassan Sawaf, Braddock Gaskill, Michael Veronis ، *Hybrid Machine Translation Applied to Media Monitoring* ، *Proceedings of the Eighth Conference of the Association for Machine Translation in the Americas, Waikiki, Hawai`i, USA, 21-25 October 2008*

