

الموسم الثقافي التاسع والثلاثين لمجمع اللغة العربية الأردني

كانون الأول ٢٠٢١

مشكلات الترجمة الآلية

أ.د. محمد زكي خضر

الخلاصة

بدأت الترجمة الآلية بشكل بدائي في منتصف القرن الماضي وتطورت تدريجياً وشملت الترجمة من اللغة العربية وإليها. تحتاج أنظمة الترجمة الآلية الحديثة إلى معاجم متعددة اللغات وذخيرة لغوية ضخمة. هذه الذخيرة تمكن الباحثين من إطلاق الوصف على خصائص اللغة من كل جوانبها بحيث تجمع غالبية الكلمات الشائعة في اللغة وغالبية التعابير اللغوية والتراكيب النحوية والصرفية. وعلى هذا فيتوقع منها أنها تمثل اللغة بشكل تقريبي. لقد أصبحت الحاجة إلى أنظمة ترجمة من العربية وإليها ماسة. وقد صدرت في الآونة الأخيرة أبحاث عديدة لتقويم أنظمة الترجمة الآلية الحديثة وذلك لمعرفة جوانب القصور فيها ومن ثم معرفة السبل اللازم سلوكها للإرتقاء بالترجمة الآلية. يحاول البحث تقصي جوانب القصور في بعض أنظمة الترجمة الآلية من جوانب نحوية وصرفية وإملائية وتركيبية وما يتعلق بالمتلازمات اللفظية.

لقد وجد أن تدخل الإنسان بعد الترجمة الآلية في الوقت الحاضر يعتبر ضرورياً لكي تكون الترجمة ناجحة وسليمة من حيث المعنى والتركييب. فالآلة تخفق في كثير من الأحيان في اختيار الكلمة المناسبة ولا تأخذ بالحسبان بعض خصوصيات ومقتضيات اللغة التي يترجم إليها والسياق الذي ترد فيه الكلمات والمفردات التي تأتي مصاحبة لها ، فتمدنا في النهاية بتراكيب تتكون من عناصر متفرقة يظهر أثرها في النص المترجم عند قراءته كأنه مجموعة من وحدات مستقل بعضها عن البعض الآخر ويعوزها الإتساق والإنسجام اللذان يمثلان شرطين ضروريين لفهم واستيعاب النص. إن مستقبل الترجمة الآلية الذي سيضمحل الترجمة الآلية نطقاً وكتابة يحتاج عدم إغفال الأبحاث المتعلقة باللغات المحلية أيضاً والتي ستكون وسيلة لخدمة اللغة العربية السليمة في المستقبل.

يعطي البحث أمثلة عملية لبعض جوانب الإخفاق في الترجمة الآلية لنصوص أدبية وعلمية وإعلامية كنماذج لمشكلات أنظمة الترجمة السائدة.

مقدمة

كان اول استخدام للحاسوب في الترجمة عام ١٩٤٧. ثم سارت بحوث الترجمة الالية قدماً بعد ذلك واجريت بنجاح أول ترجمة من اللغة الروسية الى الانكليزية عام ١٩٥٤. وفي عام ١٩٥٥ اجريت أول تجربة في الترجمة الالية من الانكليزية الى الروسية باستعمال قاموس يحوي ٢٣٠٠ كلمة. وقد استمرت الأبحاث في هذه الفترة باستعمال ما سمي الجيل الأول من برامج الترجمة الآلية حتى عام ١٩٦٦. ثم تباطأت الأبحاث بعدما تبين حجم الصعوبات الهائلة التي تعترض الترجمة الآلية وأهمل الموضوع حتى عام ١٩٧٥.

وبين عامي ١٩٧٥ و ١٩٨٥ عادت الأبحاث في هذا الحقل باستعمال الأنظمة الخبيرة وتقعيد معالجة اللغات الطبيعية واستحداث ما يمكن اعتباره الجيل الثاني من برامج الترجمة الآلية. وقد تطور في هذه الفترة البحث العلمي في معالجة اللغات الطبيعية وخاصة اللغات الأوربية واليابانية. وهذا التطور شمل البحوث المعجمية والنحو والصرف والدلالة. كما تطورت كذلك أساليب الذكاء الإصطناعي التي تستند إلى استعمال أنماط رياضية وحاسوبية تحاكي عمل الإنسان أو الكائنات الحية.

وفي عام ١٩٨٩ بدأ عهد جديد في أسس الترجمة الآلية بالإستناد إلى المعلومات الإحصائية. 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 الآلية(١).

تحتاج معالجة اللغات الطبيعية إلى ذخيرة لغوية ضخمة. هذه الذخيرة تمكن الباحثين من إطلاق الوصف على خصائص اللغة وعلى النحو فيها ودراستها تاريخياً وتغيرها مع الزمن. تجمع الذخيرة اللغوية المناسبة بين محتوياتها غالبية الكلمات الشائعة في اللغة وغالبية التعابير اللغوية والتراكيب النحوية والصرفية. وعلى هذا فيتوقع منها أن تمثل اللغة بشكل تقريبي. ونظراً لاعتماد الترجمة الآلية على الذخيرة ثنائية اللغة (أو متعددة اللغات)، فيجب أن تكون هذه الذخيرة شاملة وواسعة ودقيقة، وهو أمر في غاية الندرة. وتتوفر مثل هذه الذخيرة اليوم في اللغات العالمية الأخرى غير العربية لضخامة ما يتوفر من كتابات محوسبة وهي تمثل تلك اللغات، ومن ثم تستفيد الترجمة الآلية من مثل هذه الذخيرة. ولكن لا تزال الذخيرة اللغوية المتوفرة باللغة العربية مع ترجماتها اللغات الأخرى محدودة.

تستند عملية تهيئة الذخيرة حاسوبياً على استخراج الكلمات المكررة وغير المكررة في الذخيرة ثم بعد ذلك الحصول على قائمة بكل كلمتين متعاقبتين ثم قوائم العبارات المتقابلة بين اللغات المراد الترجمة منها وإليها.

وبالطبع فإن هذه الطريقة بأسلوبها المبسط لا تعطي ترجمة دقيقة، وهي بحاجة إلى معلومات نحوية وصرفية لزيادة الدقة. إن الخطوة التالية في معظم برمجيات الترجمة الآلية المستندة إلى الإحصاء بعد استكمال الذخيرة اللغوية المتوازية هي توليد الكلمة المقابلة بين اللغتين وحل مشكلة عدم تساوي عدد الكلمات بين اللغتين . وتعتمد دقة الترجمة الإحصائية على نوع ودقة الذخيرة المستعملة وعلى البرمجيات المستعملة فيها . ويجدر بالذكر أن تغير اللغة والمفردات المستعملة بمضي الزمن وضخامة الذخيرة المطلوبة يستوجب وجود معجم تأريخي للغة وإلى وجود قرائن في البرنامج لكي يقوم بترجمة أسماء الأعلام مثلاً بمعناها وعدم اعتبارها أسماء أعلام .. إلى غير ذلك من معلومات لزيادة دقة الترجمة(٢).

إن أحد المشاكل الرئيسية للترجمة بين اللغات المختلفة هو عدم التوافق في كثير من النواحي النحوية ومن ثم ضرورة إعادة الترتيب بين الصفات والأسماء والتوافق بين الأفعال وفواعلها والتوافق بين الضمائر وأسماء الموصول، حيث أن ترتيب الكلمات في هذه المجالات الثلاثة مختلف بين العربية والانكليزية.

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. وتحاول 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) مثلاً من الانكليزية إلى (فعل – فاعل – مفعول به) بالعربية (٣). يقوم مُحلل اللغة الطبيعية بعملية تشبه عملية إعراب الجملة نحويًا حيث يمكنه أن يفهم أجزاء الجملة ويساعد في الترجمة باستخدام الترجمة الآلية، وقد قام العديد من الباحثين بدراسة هذا النوع من المحللات مثل المحللات الصّرفية والمُحلّلات الدلالية وإعادة ترتيب الجملة (خوارزميات التصحيح) (٤).

تقويم أنظمة الترجمة الحالية

تستند أنظمة الترجمة الآلية الحديثة إلى استعمال الشبكات العصبية. وهي وسائل رياضية معقدة تستند في غالبيتها إلى تزويدها ببيانات ضخمة تقوم بالتدرب عليها ابتغاء الوصول إلى خوارزميات داخلية تستعمل لترجمة النصوص المراد ترجمتها .

انتشرت في العقود الثلاثة الأخيرة أنظمة عديدة للترجمة من العربية وإليها. ولغرض تقويم بعض هذه الأنظمة لمعرفة نقاط الضعف فيها نشرت أعداد من الأبحاث تضمنت تقويمات لبعضها. وسنستعرض بعضاً من هذه الأبحاث لغرض معرفة ما تشير إليه من نقاط الضعف في بعض تلك الأنظمة.

قامت بعض الأبحاث بمقارنة ثلاث من أنظمة الترجمة الآلية هي جوجل وبينج وجنر فوجدت أن أفضلها هو بينج حيث كانت نسبة الأخطاء ٢٦% بينما كانت النسبة لجنر ٣١% ولجوجل ٤٣%. وقد احتسبت نسبة الأخطاء وفق ما إذا استعمل النظام الكلمة الصحيحة أم لا واستعمال حروف الجر المناسبة وتسلسل الكلمات الصحيح ووضع الكلمة من نواحي الأفراد والتنثية والجمع واستعمال الضمائر المناسبة (٥).

يحتاج كثير من الطلبة لاستعمال المترجمات الآلية من اللغة الإنكليزية إلى اللغة العربية في دروسهم خاصة بسبب ضعف الطلبة باللغة الإنكليزية وكثيراً ما يلاقون صعوبة في فهم الترجمة أو حصول التباس عندما يخطئ نظام الترجمة في العبارة الصحيحة أو الكلمة الدقيقة. وقد أجريت تجارب على مدى فهم الطلبة واستفاداتهم من المترجمات الآلية والصعوبات التي يلاقونها وما تحتاج المترجمات من تحرير بعد الترجمة. وفي دراسة (٦) لتقويم استفادة الطلبة من الترجمة الآلية توصلت تلك الدراسة إلى أن الطلبة تصادفهم ترجمات فيها قصور نحوي ودلالي وعدم دقة في التعبير. وفي بعض الأحيان تكون الترجمة منحازة إلى أعراف محلية غير مناسبة لموطن في بلد يختلف العرف فيه عن البلد الذي صدرت منه الدراسة. وقد لاحظ بعض الطلبة أن الترجمة تساعدهم في وظائفهم لكن المترجمات تحتاج إلى المزيد من التطوير والتحسين.

في تلك الدراسة فإن ٦٤% من الطلبة أجابوا بأن النصوص المترجمة ليست مفهومة بسهولة لأسباب تركيبية أو نحوية أو معجمية وكان أحد أهم أسباب ضعف الترجمة الآلية هو عدم وجود مراجعة للترجمة كما يفعل المترجم من البشر لغرض تحسين التراكيب اللغوية أو زيادة الدقة أو تصليح ما فات من أخطاء.

تحديد القصور في الترجمة الآلية حالياً

قامت إحدى الدراسات في تحليل الأخطاء في الترجمة الآلية بين اللغة الإندونيسية والعربية باستعمال ترجمة جوجل. والمعروف أن مقدار الذخيرة اللغوية الثنائية بين اللغة الإندونيسية واللغة العربية محدودة، لذلك تعتمد الترجمة على اللغة الإنكليزية كوسيط بين اللغتين (٧).

وجدت الدراسة ١٦ نوعاً من الأخطاء الصرفية و ٢٢ نوعاً من الأخطاء النحوية. أما الأخطاء الإملائية فقد وجد ختان هما كتابة الياء بدل الألف المقصورة في الفعل وخطأ وضع همزة القطع في وسط الكلمة.

وفي اختيار الكلمة وجد ٧ أنواع من الأخطاء. أما أثر الأخطاء النحوية في المعنى فقد وجد أن معظمها تغير المعنى. فقد تغير المعنى في ١٥ حالة من ٢٢ خطأ نحوي. كما وجد تغير في اختيار الكلمة الخاطئ الذي يغير المعنى في ٦ كلمات من ٧ كلمات خاطئة

أما الأخطاء الصرفية فكانت : ١- إيراد المصدر بدل الفعل ٢- إيراد إسم الفاعل بدل الفعل ٣- إيراد الفعل المضارع بدل الماضي ٤- إيراد الجمع بدل المفرد ٥- زيادة ياء النسبة في إسم البلد ٦- إستخدام حرف المضارعة غير المناسب ٧- حذف تاء التأنيث الساكنة للفعل الماضي ٨- زيادة ضمير متصل لا حاجة له ٩- إيراد المصدر بدل فعل الأمر ١٠- حذف حرف المضارعة ١١- إيراد المصدر بدل فعل النهي ١٢- إيراد الفعل المضارع بدل فعل النهي ١٣- حذف ألف التثنية ١٤- صياغة المثني غير المناسبة ١٥- إيراد الفعل الماضي بدل المضارع ١٦- صياغة إسم التفضيل غير المناسب .

وقد توزعت الأخطاء النحوية على : تقديم المضاف إليه على المضاف ٢- استعمال الإسم المعرفة بدل النكرة للتمييز ٣- صيغة الجمع بدل المفرد للتمييز ٤- عدم المطابقة بين العدد والمعدود ٥- عدم المطابقة بين النعت والمنعوت ٦- تركيب الإضافة بدل النعت والمنعوت ٧- زيادة كلمة لا حاجة لها ٨- حذف مفعول فيه ٩- عدم المطابقة بين العدد الترتيبي والمعدود ١٠- حذف أداة الإستفهام ١١- حذف فعل ١٢- زيادة حرف دون حاجة ١٣- خطأ في البنية التركيبية ١٤- حذف ألف لام التعريف في إسم الموصول ١٥- إيراد تركيب النعت بدل الإضافة ١٦- استخدام إسم الموصول غير المناسب ١٧- حذف الضمير العائد للموصول ١٨- إيراد الإسم المنصوب بدل المرفوع ١٩- حذف حرف الزيادة للعبارة الإصطلاحية ٢٠- إيراد الخبر المؤنث للمبتدأ المذكر ٢١- حذف علامة التعريف للمبتدأ المذكر ٢٢- حذف كلمة (إسم).

أما الأخطاء في اختيار الكلمة فتركزت على إيراد كلمة غير مطابقة لكنها قريبة من المعنى ٢- إيراد كلمة غير مطابقة تماما لكنها بعيدة عن المعنى ٣- إيراد عبارة أو جملة ليس لها معنىً واضحًا ٤- إيراد عبارة لا علاقة لها بالمعنى ٥- إيراد عبارة لشرح معنى كلمة هي مفردة شائعة ٦- إيراد كلمة لا تناسب السياق رغم أن معناها صحيح ٧- استعمال حرف غير مناسب الوظيفة في الجملة .

وعند دراسة أثر الأخطاء في المعنى وجد أن:

أ- أثر الخطأ في المعنى كان:

١- استعمال المعرفة بدل النكرة للتمييز ٢- استعمال الجمع بدل المفرد ٣- عدم المطابقة بين العدد والمعدود ٤- عدم المطابقة بين العدد الترتيبي والمعدود ٥- خطأ البنية التركيبية لأنه يستخدم أسلوب اللغة الأخرى ٦- حذف ألف لام التعريف في إسم الموصول ٧- استخدام إسم الموصول غير المناسب.

ب- أثر الأخطاء في اختيار الكلمة في المعنى.

- ١- إيراد كلمة غير مطابقة ولكنها قريبة من المعنى المقصود ٢- إيراد كلمة غير مطابقة تمامًا لكنها بعيدة عن المعنى المقصود ٣- إيراد كلمة أو جملة ليس لها معنىً واضحًا ٤- إيراد عبارة لا علاقة لها بالكلمة ٥- إيراد عبارة تشرح كلمة هي مفردة شائعة ٦- إيراد كلمة لا تناسب السياق رغم أن معناها صحيح ٧- إيراد حرف غير مناسب للوظيفة في الجملة.

كما عالج أحد الأبحاث تغيرات المتلازمات اللفظية بين اللغات وخاصة بين العربية والإنكليزية (٨). وقد صنف البحث المتلازمات اللفظية حسب درجة تلازم الألفاظ بعضها ببعض:

- ١- المتلازمات اللفظية المفتوحة مثل "بادئ الأمر" و "أول الأمر" فالاستبدال لا يؤثر بالمعنى.
- ٢- المتلازمات اللفظية المقيدة: إحداهما ترد بالمعنى الحرفي والأخرى بالمعنى المتخصص (تقني أو مجازي أو مهمل).
- ٣- العبارات المسكوكة المجازية التي تعبر عن معنى مجازي في مجملها قد تسمح بإبدال لفظ أو أكثر بما يرادفه.
- ٤- العبارات المسكوكة المحضة وهي التي يستحيل فهم معناها بتفكيك مفرداتها لذلك لا يمكن استبدال أي من مفرداتها مثل "وضعت الحرب أوزارها".

وتتوزع المتلازمات اللفظية حسب تركيبات أجزائها بين ١- فعل وإسم ٢- صفة وإسم ٣- إسمين ٤- إسم وفعل ٥- فعل وحال أو صفة ٦- حال وصفة .

وجدت الدراسة أن الترجمة الآلية نجحت في ترجمة ١٦% من المتلازمات اللفظية الإنكليزية إلى العربية وفشلت في ٨٤% ولم تنجح على الإطلاق في ترجمة الأنواع الأنفة الذكر ١ و ٣ و ٤ و ٦ بينما نجحت في ٦٠% من النوع الخامس ونجحت بنسبة ٣٠% في النوع الثاني وقد توزع الإخفاق في أن الآلة لم تتعرف على كلا اللفظين أو أحدهما أو أنها نجحت في اختيار الألفاظ دون النجاح في حسن صياغتها أو أن الصياغة كانت جيدة لكن اختيار الألفاظ كان خاطئًا.

وقد استنتجت الدراسة أن تدخل الإنسان بعد الترجمة الآلية في الوقت الحاضر يعتبر ضروريًا خاصة في الحالات التي تكون الدقة في المحتوى أمرًا ضروريًا، فتدخل المترجم البشري بعد عملية الترجمة الآلية أمر لا مناص منه إذا كان لا بد من ترجمة ناجحة وسليمة من حيث المعنى والتركييب. فالآلة لا تأخذ بالحسبان خصوصيات ومقتضيات اللغة التي يترجم إليها والسياق الذي ترد فيه الكلمات والمفردات التي تأتي مصاحبة لها ، فتمدنا في النهاية بتراكيب تتكون من عناصر متفرقة يظهر أثرها في النص المترجم عند قراءته كأنه

مجموعة من وحدات مستقل بعضها عن البعض الآخر ويعوزها الإتساق والإنسجام اللذان يمثلان شرطين ضروريين لفهم واستيعاب النص.

إن أحد الأسباب الفنية في ضعف الترجمة الآلية في الوقت الحاضر هو قلة النصوص ثنائية اللغة بين اللغة المترجم منها والمترجم إليها. فتدريب المترجم الآلي على بيانات ضخمة يزيد من دقة الترجمة كما أن من الضروري في ترجمة النصوص المتخصصة في حقل معين توفر نصوص للتدريب مناسبة لذلك الحقل، فوجود نصوص في البيئة والزراعة والغابات والمناطق السياحية مثلا يجعل ترجمة كلمات ذات علاقة بتلك الحقول أكثر دقة (٩) .

كما أظهرت إحدى الدراسات أن الخطأ المعجمي والاختيار المعجمي غير المناسب ، وهو خطأ دلالي ، هما أكثر الأخطاء شيوعاً. وكثيراً ما يسيئ نظام الترجمة اختيار بعض الكلمات.

استنتجت الدراسة أن الترجمة غير الدقيقة يمكن الحصول عليها إذا كانت السرعة هي المقصد الرئيسي، أما إذا أريد الحصول على دقة عالية فلا تزال الترجمة البشرية هي الحل (١٠).

ومن أكثر الأبحاث تفصيلاً (١١) بحث تضمن مقارنة ثلاث برامج ترجمة هي جوجل وصخر وستران بإعطاء مجموعتين من النصوص لترجمتها من العربية إلى الإنكليزية وكذلك من الإنكليزية إلى العربية، ومن ثم البحث عن الأخطاء التي وقع فيها النوعان من الترجمة .

أعطى البحث أمثلة على الكلمة العربية التي يقابلها بالإنكليزية عدد من الكلمات المختلفة حسب السياق مثل :

بطولة : bravery, championship, heroism

إعلان : announcement, advertisement, declaration, sign

قمة : climax, peak, summit, top

يؤكد : confirm, emphasize, stress, verify

قوة : force, power, strength

مركز : centre, position, rank, status

موقع : position, rank, site, status

رئيس : major, president, primary

أسهم : arrows, shares, stocks

ومن ثم إتضح أن المشاكل الرئيسية المتعلقة بالمفردات في هذا الوضع هي أن المعجم الذي يستعمله النظام غير كافٍ، وفي حالة تعدد المعاني هذه ليس هناك ما يشير إلى معنىٍ دون آخر. وفي كل الأحوال يحتاج المعجم إلى عملية تحديث مستمرة وإضافة ما يشير لتمييز أسماء العلم للأشخاص والمدن والدول والمنظمات والجنسيات إضافة للرموز الصوتية لتمييز النطق.

أدرج الباحث في بحثه المشاكل الرئيسية المتعلقة بالقواعد والنحو في الترجمة من اللغة الإنجليزية إلى العربية هي الترتيب (حيث أن ترتيب الكلمات في الجملة العربية يختلف عن الإنكليزية) والجنس (المذكر والمؤنث وما يتبع ذلك مثل "ذلك" و"تلك" وعدم وجود جنس محايد بالعربية للجملات) وصيغة المفرد والجمع والمثنى والتحليل الخاطئ للفئات النحوية. كما أشار إلى أن هناك مشكلتان مهمتان في ترجمة جوجل هما الحذف والإضافة.

إن استعمال حروف الجر في الترجمة تشكل معضلة. فهي تحتاج إعادة تموضع وأن لها معانٍ مختلفة حسب الاسم أو الفعل المجاور لها. يبدو أن أفضل طريقة لمعالجة حروف الجر هي إنشاء صلة بين الأفعال والأسماء مع حروف الجر بحيث لا يتم التعامل مع حروف الجر بشكل منفصل قدر الإمكان. في اللغة العربية يجب أن تحتوي الأسماء العامة والعديد من أسماء الأماكن والعناوين على ألف لام التعريف. فكثيراً ما تخطئ أنظمة الترجمة في ذلك (١٢).

هناك صعوبة في إدراك الآلة التمييز الدقيق بين الكلمات من نوع واحد مثل "هنا" و"هناك" وكذلك الضمائر المستترة وآثار اللواحق في الكلمات المركبة. إن توفر معرفة لغوية واسعة للمترجم الآلي يساعد في رفع مستوى الترجمة ودقتها (١٣).

من المشاكل الخاصة باللغة العربية تعدد معاني الحروف. فعلى سبيل المثال فإن الحرف "على" يمكن أن يستعمل للإستعلاء الحقيقي – المجازي – زائدة – بمعنى من – بمعنى في – بمعنى مع – السببية للتعليل – بمعنى عن – بمعنى الباء – الاستدراك ومثل ذلك تتعدد استعمالات بقية الحروف.

كما أن قلة استعمال العلامات أو الخطأ فيها يزيد الصعوبة في الفهم الصحيح للنصوص ومن جهة أخرى عدم وجود اتفاق شامل لكيفية استخدامها في النصوص المترجمة مما يعطي القائمين على برمجة نظام الترجمة الاجتهادات في استعمالات مختلفة للعلامات. كما أن عدم التشكيل ينتج عنه الكثير من الأخطاء البشرية، فكيف بالآلة.

تعاني معظم المعاجم العربية من الضعف في تحديد المعاني ومن ضعف الأمثلة التي تزيل الغموض في استعمالات الكلمة الواحدة وفي التصاق حرف الجر بالفعل وتغيرات المعاني نتيجة تغير حرف الجر.

أما مشكلة عودة الضمير فهي تسبب التباسًا حتى للقارئ من البشر فكيف هي الحال بالنسبة للآلة. إن المشاكل المتعلقة بتوليد المثنى وجمع المذكر السالم وجمع المؤنث السالم ونون النسوة ونون التوكيد لا يمكن حلّها إلا من خلال معرفة القرائن المناسبة لتوليدها.

أما النصوص البلاغية فهي صعبة جدًا على الآلة وهي مرحلة متقدمة في الاستفادة من الترجمة الآلية ومن المبكر الخوض فيها ربما تحتاج عقدًا أو أكثر من الزمان حتى يمكن توليد نصوص ذات مستوى بلاغي مقبول باللغة العربية.

مشاكل اللهجات المحلية

إهتمت بعض الأبحاث بدراسة اللهجات العربية المحلية بهدف الترجمة من هذه اللهجات إلى اللغة الإنكليزية. لكن ما يهمننا في دراسة اللهجات المحلية هو تحويل الكلام المنطوق باللهجات المحلية إلى كلام مكتوب بلغة عربية سليمة. وقد ينطبق على هذه العملية من بعض الجوانب ما ينطبق على الترجمة الآلية. تشكو مثل هذه المحاولات من شح الذخيرة اللغوية بين اللهجات المحلية واللغة العربية السليمة.

قام بعض الباحثين بدراسة التحويل بين اللهجات المحلية واللغة العربية السليمة في مصر وبلاد الشام والخليج وشمال أفريقيا (١٤). وتصادف مثل هذه الدراسات الإختلاف الشديد في اللهجات بين قطر وآخر بل بين المدن المختلفة داخل القطر الواحد. وإذا كان الهدف هو نظام ترجمة آلية يستقبل الكلام المنطوق بلهجة عامية فكيف يكون التصرف إذا غيّر المتكلم نفسه من لهجة إلى أخرى؟. لذلك لا تبدو المهمة سهلة وميسورة ، لكن المستقبل لا بد أن يعالج مثل تلك الصعوبات.

وفي بحث آخر يتعلق باللهجات المحلية أشير كذلك إلى نقص موارد البيانات. فبعد استعمال ذخيرة مكونة من ١,١ مليون كلمة بلهجة مصرية و ٣٨٠ ألف كلمة بلهجة شامية اختيرت من مواقع على الشابك، ومع هذا وجد أن ذلك غير كافٍ (١٥).

أمثلة

أختير في هذا البحث ثلاث نصوص لتجربة الترجمة على ثلاث أنظمة ترجمة هي جوجل وبينج وسستران للترجمة من الإنكليزية إلى العربية. النص الأول هو نص لخبر صحفي عن رحلة فضائية والثاني لمقطوعة أدبية عن القمر والثالث هو نص علمي عن الجاذبية الأرضية.

يبين الجدول (١) بعض الأخطاء التي ارتكبت في الترجمة وبعض الملاحظات عنها. ويبين الجدول الركافة في بعض الكلمات والخطأ الواضح في ترجمة البعض الآخر.

جدول (١) الأخطاء في ترجمة معاني كلمات

الكلمة	ترجمتها	برنامج الترجمة	النص الوارد فيه	الملاحظات
rode out	ركب	جوجل	الإعلام	ضعيفة
=	خرج	سستران وبنج		بعيدة
The quartet	الرباعي	جوجل	الإعلام	ضعيفة
=	الرباعية	بنج		ضعيفة
=	المجموعة الرباعية	سستران		أفضل
no clue	بدون أدنى فكرة	جوجل	الأدب	ركيكة
=	لا يزال أي فكرة	بنج		خاطئة
=	لا يزالون لا يعرفون	سستران		ضعيفة
how to harness electricity	كيفية تسخير الكهرباء	جوجل وبنج	الأدب	خاطئة
=	كيف يستعملون الكهرباء	سستران		لا بأس
create entertainment	خلق الترفيه	جوجل وبنج	الأدب	ضعيفة
=	خلق التسلية	سستران		ضعيفة
most loyal friend	أكثر الأصدقاء ولاءً	جوجل	الأدب	ركيكة
=	الصديق الأكثر ولاءً	بنج		ركيكة
=	أكثر صديق	سستران		خاطئة
untinged by artificial light sources	لا تتأثر بمصادر الضوء الاصطناعي	جوجل	الأدب	لا بأس

خاطئة		سستران	لا تلدغ بمصادر النور الاصطناعية	=
دمج الإنكليزية غير مناسب		بنج	untinged من مصادر الضوء الاصطناعي	=
خاطئ	الأدب	جوجل وبينج	القمر الصناعي الطبيعي الوحيد للأرض	only natural satellite of earth
خاطئ		سستران	القمر الطبيعي الوحيد على وجه الأرض	=
ضعيفة	الأدب	جوجل	اللوحة اللامحدودة	boundless canvas
ضعيفة		بنج	اللوحة التي لا حدود لها	=
خاطئة		سستران	القماش الذي لا يحد	=
لا بأس	الأدب	جوجل وسستران	يومض	blinking
غير مناسبة		بنج	يرمش	=
الصحيح هو الروحانية	الأدب	جوجل وسستران	التصوف	mysticism
خاطئة		بنج	الصوف	=
ضعيفة	العلمية	جوجل	تجعيد	crumple
خاطئة		بنج	تكوم	=
خاطئة		سستران	تكسير	=
ضعيفة	العلمية	جوجل وسستران	كرة مضغوطة	compact ball
خاطئة		بنج	كرة مدمجة	=
ضعيفة	العلمية	جوجل	تضرب الأرض	hit the floor
لا بأس		سستران	تصطدم بالأرض	=
خاطئة		بنج	ضرب الكلمة	=

ضعيفة	العلمية	جوجل	لكن أيضًا على مساحته	but also area
خاطئة		بنج	لكن أيضا المنطقة	=
ضعيفة		سستران	بل على مساحته أيضا	=
خاطئة	الأدب	جوجل وبنج	القمر - القمر الصناعي الطبيعي الوحيد للأرض	the Moon – the one and only natural satellite
خاطئة		سستران	القمر القمر القمر وهو القمر الطبيعي الوحيد على وجه الأرض	=
لا بأس	الأدب	جوجل	أبسط وأشهر	most common
لا بأس		بنج	أبسط وأكثر الأمثلة شيوعا	=
ضعيفة		سستران	أبسط وأشيع	=
ضعيفة	الأدب	جوجل	تحمل أكبر قدر من المشاعر والرضا	has shouldered the most emotion and favor
خاطئة		بنج	تحمل ... أكبر قدر من العاطفة والتفضيل	=
خاطئة		سستران	يحمل القدر الأعظم من المشاعر والنعمة	=

المناقشة والإستنتاجات

تكمن الصعوبات في الترجمة الآلية في القصور في فهم النص من اللغة المترجم منها أو القصور في توليد النص الدقيق في اللغة المترجم إليها. وقد حاولت هذه الورقة إعطاء فكرة عن الواقع الحالي لتقنيات الترجمة الآلية والمشاكل التي تعاني منها في الوقت الحاضر. إن مشاكل الترجمة الآلية يمكن أن تتعلق بفهم الآلة للغة المترجم منها وإمكانية الآلة للتعبير عن هذا الفهم باللغة المترجم إليها. لقد لاقت الترجمة من اللغة العربية إلى

لغات أخرى وخاصة إلى اللغة الإنكليزية أبحاثاً كثيرة بعكس الحال فيما يخص الترجمة من اللغات الأخرى إلى اللغة العربية ، وربما يعود السبب إلى الإهتمام العالمي للتعرف على ما ينشر باللغة العربية بينما الترجمة إلى اللغة العربية تكون فائدتها بالدرجة الأولى إلى المتكلمين باللغة العربية. وهذا يضع الواجب بالدرجة الأولى على الباحثين العرب لتدارك هذا الواقع ذلك لأن الترجمة إلى اللغة العربية ذات أهمية بالغة للعرب اليوم. لقد وصل تحليل النصوص باللغات العالمية (الأوروبية والصينية واليابانية والكورية) إلى حد كبير من التقدم كما أن الذخائر اللغوية التي تحتاجها الترجمة الآلية بين تلك اللغات متوفرة بشكل كبير، مما يوفر المادة الخام الرئيسية لأنظمة الترجمة الآلية بينها والتي ترفع من دقتها لحد كبير. أما بالنسبة للغة العربية، فإن الذخائر اللغوية متعددة اللغات (التي تكون اللغة العربية واحدة منها) لا تزال غير كافية للوصول إلى الدقة المطلوبة. حاولت هذه الورقة التركيز على الترجمة الآلية إلى اللغة العربية، لذلك فإن عملية فهم الآلة للغة المترجم منها يفترض أن يكون دقيقاً بدرجة لا بأس بها، نظراً للتقدم الكبير في معالجة نصوص اللغة الإنكليزية. أما دقة تكوين النصوص العربية المترجمة من الإنكليزية فهي ما يهمنى هنا.

من الأمثلة المعروضة في هذه الورقة وجد أن أفضل ترجمة هي للنص الإعلامي نظراً لتوفر كمية هائلة من النصوص الإعلامية على الشابكة يليها النصوص العلمية ثم النصوص الأدبية قليلة التوافر.

ومن الواضح أن التنقيح بعد الترجمة الآلية ضروري إذا ما أريد الوصول إلى ترجمة دقيقة وواضحة.

المراجع

١- محمد زكي خضر - اللغة العربية والترجمة الآلية - المشاكل والحلول ، مؤتمر التعريب الحادي عشر -

المنظمة العربية للتربية والثقافة والعلوم - عمان ١٢-١٦/١٠/٢٠٠٨

٢- محمد زكي خضر - الترجمة الآلية من العربية وإليها - الحرف العربي والتقنية - أبحاث في حوسبة اللغة

العربية - مركز الملك عبدالله بن عبد العزيز الدولي لخدمة اللغة العربية ٢٠١٥

٣ - محمد زكي خضر - الترجمة الآلية إلى اللغة العربية ضرورة ملحة - المؤتمر الدولي السابع للغة العربية

- ١٧-٢١/٤/٢٠١٨ دبي

٤- محمد زكي خضر - الترجمة الآلية في خدمة اللغة العربية - الموسم الثقافي السادس والثلاثين

لمجمع اللغة العربية - كانون الأول ٢٠١٨

5-

Mohammed Abdulmalik Ali, An Evaluation of Online Machine Translation of English into Arabic Texts, Open Journal of Modern Linguistics, 2020, 10, 524-548

6-

Amin Ali AlMubark, The Problems of Machine Translation From Arabic Into English language Faced by Saudi University Student Of Translation At The Faculty-
- Of Arts, Jazan University Saudi Arabia University Saudi Arabia

IOSR Journal of Humanities and Social Science (IOSR-JHSS) Volume 21, Issue 4,
Ver. II (Apr. 2016) PP 55-66

٧- سهيا حسين - تحليل الأخطاء في الترجمة الآلية - دراسة عن ترجمة اللغة الأندونيسية إلى اللغة العربية
من محرك جوجل - رسالة ماجستير - جامعة مولانا مالك إبراهيم الإسلامية الحكومية بمالانج -
٢٠١٤

٨- أمانة فاطمة الزهراء طالبي - إشكالية حدود الترجمة الآلية - ترجمة نظام "سيستران" للمتلازمات
اللفظية - رسالة ماجستير - قسم الترجمة - جامعة منتوري - قسنطينة - الجزائر - ٢٠٠٧/٢٠٠٨

9- Wael Abid, Younes Bensouda Mourri, Improving English to Arabic Machine Translation - 32nd Conference on Neural Information Processing Systems (NIPS 2018), Montréal, Canada.

10- Noureldin Mohamed Abdelaal, Theory and Practice in Language Studies, Machine Translation: The Case of Arabic- English Translation of News Texts Vol. 10, No. 4, pp. 408-418, April 2020

11- Sattar Izwaini, Problems of Arabic Machine Translation: evaluation of three systems, Proceedings of the International Conference on the Challenge

of Arabic for NLP/MT. The British Computer Society (BSC), London 23 Oct. 2006. 118-148.

- 12- Eltayeb Abuelyaman, Wafaa Mukhtar, Limia Rahmatallah, Mona Elagabani, Machine Translation of Arabic Language: Challenges and Keys, 2014 Fifth International Conference on Intelligent Systems, Modelling and Simulation
- 13- Mouiad Alawneh, Nazlia Omar and Tengku Mohd Sembok, Machine Translation from English to Arabic, 2011 International Conference on Biomedical Engineering and Technology IPCBEE vol.11 (2011) © (2011) IACSIT Press, Singapore
- 14- Salima Harrat, Karima Meftouh, Kamel Smaïli, Machine translation for Arabic dialects (survey), Information Processing and Management, Elsevier, 2017, 56 (2), pp.262-273.
- 15- Rabih Zbib, Erika Malchiodi, Jacob Devlin, David Stallard, Spyros Matsoukas, Richard Schwartz, John Makhoul, Omar F. Zaidan, and Chris Callison-Burch, Machine translation of arabic dialects. The 2012 Conference of the North American Chapter of the Association for Computational Linguistics, Montreal, June.

ملحق ١ مثال لخبر إعلامي

Four amateur astronauts have blasted off from Florida on their private mission to orbit.

The Inspiration4 crew, comprising one billionaire and three "ordinary citizens", rode out of the Kennedy Space Center in a Dragon capsule provided by the SpaceX rocket company.

The quartet will spend the next three days circling the Earth.

جوجل

انطلق أربعة رواد فضاء هواة من فلوريدا في مهمتهم الخاصة إلى المدار. ركب طاقم Inspiration4 ، الذي يتألف من ملياردير وثلاثة "مواطنين عاديين" ، من مركز كينيدي للفضاء في كبسولة Dragon التي قدمتها شركة SpaceX للصواريخ. ستقضي الرباعية الأيام الثلاثة القادمة في الدوران حول الأرض. (الترجمة احتفظت بكلمات إنكليزية لأسماء الأعيان – المدار بالعربية تحتاج تعريف أي المدار حول الأرض)

بينج

انطلق أربعة رواد فضاء هواة من فلوريدا في مهمتهم الخاصة إلى المدار. خرج طاقم "إمالين ٤" ، الذي يضم مليارديرا وثلاثة "مواطنين عاديين" ، من مركز كينيدي للفضاء في كبسولة تتين قدمتها شركة سبيس إكس للصواريخ. الرباعي سيقضي الأيام الثلاثة القادمة حول الأرض (ترجم أحد أسماء العلم)

سستران

إنطلق أربعة رواد فضاء من الهواة من فلوريدا في مهمتهم الخاصة إلى المدار. خرج طاقم "إليهام 4" المكون من ملياردير وثلاثة "مواطنين عاديين" من مركز كينيدي الفضائي في كبسولة تتين وفرتها شركة "سبيس إكس" الصاروخية. وسوف تقضي المجموعة الرباعية الأيام الثلاثة المقبلة وهي تدور حول الأرض. (ترجمت أسماء الأعلام جزئيا)

(حروف الجر : من – في – إلى – لـ - استعملت بشكل مناسب في الترجمات الثلاثة)

(استعمال الجمل الفعلية بدل الجملة الاسمية جرى بشكل مناسب)

ملحق ٢ مثال لنص أدبي

Back in the days when people still had no clue as how to harness electricity, not to mention utilizing its power to create entertainment, the night sky, untinged by

artificial light sources, was the most loyal friend one could have after sunset. No wonder all of our ancestors have projected their abundant imaginations and creativity onto this boundless canvas with stars scattering across it despite the differences of civilizations.

And among all of this diamond dust blinking to us from the distance, the Moon – the one and only natural satellite of the earth – has shouldered the most emotion and favor from mankind through the ages. Due to its relatively gentle and soothing beam and its subtle connection with the tide, the Moon has long been considered as a symbol of mysticism and femininity almost ubiquitously around the world.

جوجل

بالعودة إلى الأيام التي كان الناس لا يزالون فيها بدون أدنى فكرة عن كيفية تسخير الكهرباء ، ناهيك عن استخدام قوتها لخلق الترفيه ، كانت سماء الليل ، التي لا تتأثر بمصادر الضوء الاصطناعي ، هي أكثر الأصدقاء ولاءً الذي يمكن للمرء الحصول عليه بعد غروب الشمس. لا عجب أن جميع أسلافنا قد عرضوا خيالهم الوافر وإبداعهم على هذه اللوحة اللامحدودة مع النجوم المنتشرة عبرها على الرغم من اختلافات الحضارات. ومن بين كل هذا الغبار الماسي الذي يومض لنا من بعيد ، فإن القمر - القمر الصناعي الطبيعي الوحيد للأرض - قد تحمل أكبر قدر من المشاعر والرضا من البشرية عبر العصور. نظرًا لشعاعه اللطيف والمريح نسبيًا واتصاله الدقيق بالمد والجزر ، لطالما اعتبر القمر رمزًا للتصوف والأنوثة في كل مكان تقريبًا حول العالم.

بينج

مرة أخرى في الأيام التي كان الناس لا يزال أي فكرة عن كيفية تسخير الكهرباء، ناهيك عن استخدام قوتها لخلق الترفيه، والسماء ليلاً، untinged من مصادر الضوء الاصطناعي، وكان الصديق الأكثر ولاء يمكن للمرء أن يكون بعد غروب الشمس. لا عجب أن جميع أسلافنا قد عرضوا خيالهم الوفير وإبداعهم على هذه اللوحة التي لا حدود لها مع النجوم المتناثرة عبرها على الرغم من اختلافات الحضارات .

ومن بين كل هذا الغبار الماسي الذي يرمش لنا من مسافة بعيدة، تحمل القمر - القمر الصناعي الطبيعي الوحيد للأرض - أكبر قدر من العاطفة والتفضيل من البشرية عبر العصور. نظرا لشعاعه اللطيف والمهدئ نسبيا

واتصاله الدقيق بالمد والجزر ، فقد اعتبر القمر منذ فترة طويلة رمزا للصوف والأوثة في كل مكان تقريبا في جميع أنحاء العالم.

(العبارة " لا يزال أي فكرة عن كيفية تسخير الكهرباء " خاطئة تماماً وواضح فيها الخلط بين المفرد والجمع حيث كان يجب أن يقال: الناس لا يزالون)

(استعمال حروف الجر كان يمكن أن يكون افضل : " في كل مكان تقريبا في جميع أنحاء العالم " كان يمكن أن تكون " في كل مكان من العالم تقريبا") كما استعملت كلمات خاطئة مثل "القمر الصناعي الطبيعي" بدل أن يقال "التابع" .

سستران

ففي الايام التي كان الناس لا يزالون لا يعرفون كيف يستعملون الكهرباء، بالإضافة إلى استعمال قوتها لخلق التسلية، كانت سماء الليل، التي لا تلدغ بمصادر النور الاصطناعية، أكثر صديق ولاء يمكن ان يملكه المرء بعد غروب الشمس.

لا عجب أن جميع أجدادنا أقدموا على تصوراتهم وإبداعاتهم الوفيرة على هذا القماش الذي لا يحد، مع النجوم المبعثرة عليه برغم أختلاف الحضارات

ومن بين كل هذا الغبار الماسي الذي يومض لنا من بعيد، كان القمر القمر القمر وهو القمر الطبيعي الوحيد على وجه الأرض يحمل القدر الأعظم من المشاعر والنعمة من البشر على مر العصور .وبسبب شعاعه اللين والمهدئ نسبيا وصلته الدقيقة بالمد، اعتبر القمر منذ زمن طويل رمزا إلى التصوف والأوثة في كل مكان تقريبا حول العالم .

ملحق ٣ مثال لنص علمي

The motion of falling objects is the simplest and most common example of motion with changing velocity. If a coin and a piece of paper are simultaneously dropped side by side, the paper takes much longer to hit the ground. However, if you crumple the paper into a compact ball and drop the items again, it will look like both the coin and the paper hit the floor simultaneously. This is because the amount of force acting

on an object is a function of not only its mass, but also area. Free fall is the motion of a body where its weight is the only force acting on an object.

جوجل

حركة الأجسام الساقطة هي أبسط وأشهر مثال للحركة ذات السرعة المتغيرة. إذا تم إسقاط عملة وقطعة من الورق جنبًا إلى جنب في نفس الوقت ، فإن الورقة تستغرق وقتًا أطول لتصل إلى الأرض. ومع ذلك ، إذا قمت بتجعيد الورق إلى كرة مضغوطة وإسقاط العناصر مرة أخرى ، فسيبدو كل من العملة المعدنية والورقة تضرب الأرض في وقت واحد. هذا لأن مقدار القوة المؤثرة على جسم ما لا يقتصر فقط على كتلته ، ولكن أيضًا على مساحته. السقوط الحر هو حركة الجسم حيث يكون وزنه هو القوة الوحيدة المؤثرة على الجسم.

بينج

حركة الأجسام الساقطة هي أبسط وأكثر الأمثلة شيوعًا للحركة مع تغير السرعة. إذا تم إسقاط عملة معدنية وقطعة من الورق في وقت واحد جنبًا إلى جنب ، فإن الورق يستغرق وقتًا أطول بكثير لتصل إلى الأرض. ومع ذلك ، إذا كنت تكوم ورقة في كرة مدمجة وإسقاط العناصر مرة أخرى ، وسوف تبدو وكأنها كل من العملة والورق تضرب الكلمة في وقت واحد. وذلك لأن مقدار القوة التي تعمل على جسم ما هو دالة ليس فقط كتلته ، ولكن أيضًا المنطقة. السقوط الحر هو حركة الجسم حيث وزنه هو القوة الوحيدة التي تعمل على كائن. (هناك أخطاء واضحة في الذكير والتأنيث وفي سلاسة الصياغة والتعبير)

سستران

حركة الأجسام الساقطة هي أبسط وأشيع مثال للحركة ذات السرعة المتغيرة. وإذا سقطت قطعة نقدية وقطعة من الورق في آن واحد جنبًا إلى جنب ، يستغرق وصول الورقة إلى الأرض وقتًا أطول بكثير. ولكن إذا قمت بتكسير الورقة في كرة مضغوطة وقمت بإسقاط العناصر مرة أخرى ، فستبدو وكأن كل من العملة المعدنية والورقة تصطدم بالأرض في وقت واحد. وذلك لأن كمية القوة التي تؤثر على جسم ما ليست دالة على

كتلته فحسب، بل على مساحته أيضا. السقوط الحر هو حركة الجسم حيث يكون وزنه هو القوة الوحيدة المؤثرة على جسم ما.