نحو ترجمة آلية بسمات بشرية للنصوص المتخصصة من اللغة الإنجليزية إلى العربية :دراسة مقارِنة

Vers une traduction automatique d›inspiration humaine des textes spécialisés de l›anglais vers l›arabe. Étude comparative

Towards Human Inspired Machine Translation of English to Arabic Specialized texts. Comparative Study

مختار بن ونان هاجرBenounane Hadjer
مخبر اللغة والتواصل -جامعة أحمد بن بلة - وهران1
ناصر جيلالي Nacer Djilali
مخبر اللغة والتواصل - جامعة أحمد بن بلة- وهران1

مقدمة

أدت الثورة التقنية إلى تطوير حقول المعرفة بشكل عام وميدان الترجمة بشكل خاص، وأخذ اعتمادُ المترجم على الآلة يزداد يوما بعد يوم، بل صار إدخالها في عمله من مظاهر التقدم، بينما يعتبر غير ذلك من مظاهر التخلف وسوء التكيف. ومع أن الإنسان ما يزال يعتمد على قدراته الذاتية في إنجاز مهام الترجمة بين اللغات المختلفة، إلا أن تطوير امكانيات الحاسوب من حيث سعة التخزين وسرعة المعالجة شجّع على إدخال أدوات الترجمة الآلية إلى منصة عمل المترجم. غير أن التعامل مع اللغة الطبيعية أمر يعتمد على الملكة العقلية للبشر، وهذه الأخيرة ليست عملا آليا كما هو الشأن في الأمور الأخرى كالتصنيع وتسيير المركبات، بل تقوم على ملكات التفكير والتحليل والبحث والاستنباط.

ومن هذا المنطلق، نشأت فكرة محاكاة بنية المخ البشري في الترجمة الآلية بما يحتوي عليه من مليارات العصبونات المتشابكة في شكل شبكات اصطناعية داخل الحاسوب في تسعينيات القرن الماضي، وعادت لتظهر من جديد مؤخرا لتهدّد مكانة الترجمة الآلية الإحصائية إلى الزوال، ونجحت في ذلك إلى حد كبير، حيث أنه لم يمر على عمر الترجمة الآلية العصبية الآن سوى خمسة أعوام ونجدها قد فاقت المناهج السابقة من حيث الأداء بشكل ملحوظ. وعليه، يسعى هذا البحث لدراسة حيثيات استبدال المنهج الإحصائي بالنموذج العصبي، ومدى تأثير هذا التغير على ترجمة النص المتخصص من اللغة الإنجليزية إلى العربية.

وفي سعينا لدراسة هذه الإشكالية، قسمنا بحثنا إلى القسم النظري الذي يشمل مدخلا للنص المتخصص وخصائصه، ثم لمحة تاريخية حول منهج الترجمة الآلية الإحصائية والعصبية ومفهومهما وأوجه الشبه والاختلاف بينهما، لننتقل بعد ذلك إلى الدراسة المقارنة بين أداة Reverso وGoogle Translate وتقييم نتاج الترجمة الآلية لنص متخصص من اللغة الإنجليزية إلى العربية وتحليل الأخطاء الواردة فيه، للخروج بجملة من التوصيات المدرجة في خاتمة البحث.

1. النص المتخصص في حدود المفهوم

يعتمد العلم في أي مجال من المجالات على تدوين المعلومات والمعارف وتخزينها في نصوص وصور وأشكال بهدف تسهيل دراستها ومشاركتها. ويُعد النص المتخصص المادة الأولية المسؤولة عن توثيق الأفكار العلمية وتنظيمها، إذ لا يمكن لأي علم من العلوم أن ينشأ أو يتطور بعيدا عن الخطاب العلمي كونه الوسيلة الأكثر فعالية واستدامة في التعبير وتحقيق التواصل. ويقصد بالخطاب العلمي في هذا السياق لغة التخصص التي تقوم على جملة من الوسائط اللغوية التي تبرز قيمتها ووظيفتها في سياق تواصلي ذي طبيعة علمية بحتة (حجازي، محمود، 1995 :9).

2. الترجمة الآلية الإحصائية والعصبية : نظرة تاريخية

في سنة 1990، وبعد فشل الترجمة القائمة على القواعد وعلى الأمثلة في تجسيد فكرة الترجمة الآلية كما كان مخططا لها، وضعت شركة IBM منظومة للترجمة من الفرنسية إلى الإنجليزية اعتمدت فيها على نصوص مترجمة مسبقا لاستنباط احتمالات وجود كلمة إنجليزية تصلح لأن تكون ترجمة صحيحة للكلمة الفرنسية (Brown ,.al et 1990 :79-80)، وهي الطريقة ذاتها التي تعتمدها برامج الترجمة الآلية الإحصائية إلى يومنا هذا ولكن بشكل أكثر تعقيدا. غير أن هذا المنهج -ورغم أدائه المقبول في الترجمة-لم يسلم من الانتقاد بسبب المشاكل اللغوية التي يمكن أن تظهر في حال ما إذا كانت الذخيرة اللغوية محدودة المصادر، وهذا ما دفع الوسط البحثي للتفكير في استبداله بمنهج يحاكي طريقة عمل الإنسان من خلال تصميم شبكة عصبية اصطناعية قابلة للتعلم والتدرب.

وعلى الرغم من أن الشبكات العصبية لم تُعرف سوى منذ وقت قصير إلا أن تطويرها امتد لسنوات وبالتحديد في سنة 1997 حين قدم Castaño وطاقمه البحثي (Castaño 1997 : 161) اقتراحات أولية لاستعمال نماذج قائمة على شبكات عصبية اصطناعية في ترجمة الجمل آليا، غير أن هذه التجارب لم تحض بالاهتمام الكافي بسبب أن امكانيات

الحاسوب التقنية والتخزينية لم تكن بالتطور الذي يمكّنها من تطبيق هذه الفكرة على أرض الواقع. إلا أن الأبحاث في هذا الصدد تواصلت بفضل تجارب Scwenk في 2007 لدمج نموذج الشبكة العصبية الاصطناعية في نظم الترجمة الآلية الإحصائية (Koehn 2010 : 9)، وأبحاث Blunsom وKalchbrenner سنة 2013 لاقتراح منهج الطرفين للتشفيروفك التشفير(end to end encoder-decoder approach) على برامج الترجمة الآلية. Alrajeh, 2018 :1))

وعلى إثر هذه الأبحاث عادت فكرة الترجمة الآلية العصبية لتَرى النور من جديد وبشكل أكثر جدية سنة 2014 لتوَفُّر الشرطين الأساسيين لتحويل هذه المقاربة إلى حقيقة وهما « وجود قدرة حاسوبية عالية بقدر كاف، وتوفر كميات كبيرة من البيانات لتدريب الشبكات العصبية بواسطتها » (كروجر رالف، ت. فؤاد لبنى 2018 (، حيث تم إطلاق أول نموذج عصبي قائم على ترجمة مقاطع قصيرة (sequence to sequence model) والذي نجح في تقديم نتائج أولية مرضية، خاصة بعد تدعيم الشبكة العصبية الاصطناعية بآلية الانتباه سنة 2015 والمصادقة على إطلاق أول برنامج ترجمة آلية عصبية تجريبي في مؤتمر الترجمة الآلية الدولي من نفس السنة، وهذا ما شجع العديد من المؤسسات الكبرى أمثال Google وMicrosoftو Systran لتحويل مسار برامجها من المنهج الإحصائي إلى النموذج العصبي بداية سنة 2016 وإضافة المزيد من الأزواج اللغوية وتغذية قواعد بياناتها بشكل متواصل للحفاظ على تنافسيتها في سوق الترجمة العالمي. وفي الوقت الذي نحرر فيه بحثنا هذا، تتواصل مشاريع تطوير أداوت الترجمة الآلية العصبية وتحسين أدائها لتصبح في المستقبل جاهزة لتجاوز المشاكل اللغوية التي عجزت المناهج السابقة عن حلها.

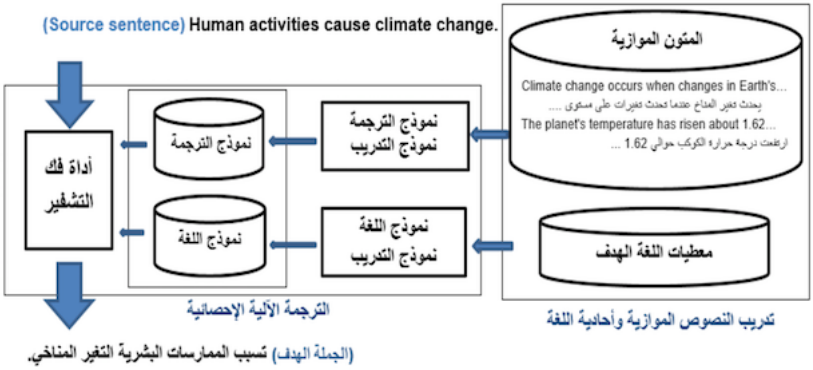
وأكثر ما يميز الخطاب العلمي هو امتلاكه لمعجم مصطلحي متخصص يميزه عن باقي الخطابات الأخرى، ويعرف « المصطلح على أنه كلمة أو مجموعة من الكلمات من لغة متخصصة، موروثا أو مقترضا ويستخدم للتعبير بدقة عن المفاهيم وليدل على أشياء مادية محددة » (حجازي، محمود 1995 : 11)، والمصطلح نتاج اتفاق بين أهل الاختصاص على تسمية مادة أو ظاهرة معينة بمصطلح ما، فعلى عكس اللفظ العام الذي يقبل تعدد الدلالات والتأويلات حسب السياق يحمل المصطلح معنى واحدا وثابتا في حدود حقل معرفي معين.

وتستدعي عملية وضع المصطلحات الكثير من الدقة والصرامة وتستوجب من صاحبها التخصص في الميدان العلمي والتعمق في خباياه وكذا إتقان اللغة إتقانا تاما حتى يتمكن من اختيار المصطلح السليم وتحديد مفهومه بدقة، حيث تفرض المنظمة العالمية للتقييس

(ISO) على اللجان المختصة في وضع المصطلحات احترام مبدأ الاقتصاد اللغوي باختيار ألفاظ قصيرة وسهلة التلفظ ومراعاة قواعد اللغة الصوتية والصرفية وضرورة صياغة مصطلح قابل للاشتقاق والتأكد من وضوح معناه عند وضعه في سياق منفصل.

3. الترجمة الآلية الإحصائية : قراءةٌ في المفهوم والتحديات

تنطلق مقاربة الترجمة الآلية الإحصائية من مبدأ أن الترجمة مهارة بشرية كونها « تستعمل متونا لغوية أحادية وثنائية اللغة مترجمة من قِبل مترجمين محترفين، إذ تُستعمل المتون أحادية اللغة لتدريب نماذج اللغة والمتون ثنائية اللغة لتدريب نماذج الترجمة » (بن تركي وفاء، سمار نصر الدين 2013 : 117)، أي أنها تعتمد على النصوص أحادية اللغة لتحليل قواعد النص المدخل والنصوص ثنائية اللغة لترجمة النص الأصل إلى اللغة الهدف.



شكل (01) : طريقة عمل أداة الترجمة الآلية الإحصائية

ولقد بدأ استعمال منهج الترجمة الآلية الإحصائية في Google سنة 2006 واستمرت الشركة في تبني هذا المنهج إلى غاية سبتمبر 2016، على عكس Babylon Translator و Reverso اللذان لا يزالان يعتمدان على منهج الترجمة الآلية الإحصائية إلى يومنا هذا. أما Moses فهو في الأساس أداة فك التشفير (encoding toolkit) مفتوحة المصدر تمكّن المترجمين من تغذية قاعدة بيانات البرنامج بمصادرهم اللغوية من متون أحادية ومتعددة اللغات ثم ترجمة النصوص على نفس الأداة من خلال تحديد الجملة ذات أعلى احتمال. كما يسمح البرنامج بإظهار جميع الجمل المرشحة لتقديم ترجمة سليمة بهدف توسيع خيارات المترجم وتمكينه من الحصول على ترجمة أكثر دقة ومقاربة لسياق النص (بن تركي

وفاء، سمار نصر الدين 2013 : 121). ويقدم البرنامج أيضا جميع المعلومات التي أدت به لاختيار جملة معينة من بين آلاف الاحتمالات.

4. مآخذ الترجمة الآلية الإحصائية

إن أداء نظم الترجمة الآلية الإحصائية مرهون بنوعية وحجم النصوص المخزنة في قاعدة بياناتها، وفي حال عدم وفرة وتنوع الذخيرة اللغوية، فسيُنتج البرنامج نصا مترجما ذا مشاكل لغوية متعلقة ب :

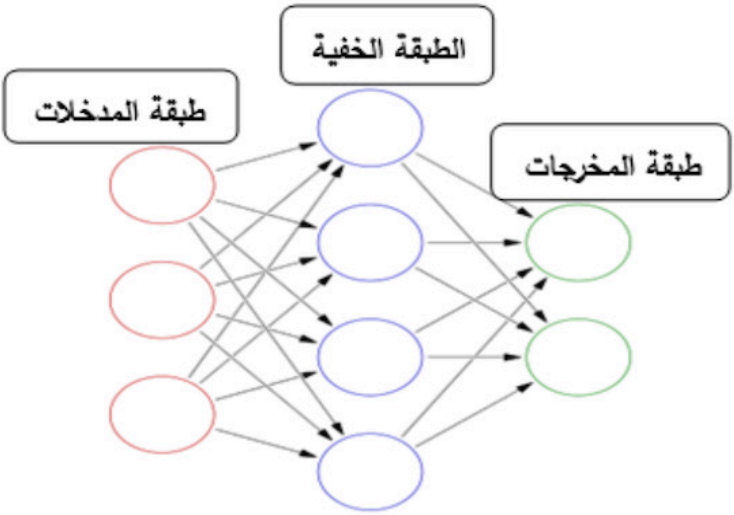
- محاذاة الجمل sentence alignment)) : من المعلوم أن لكل لغة منطقها وطريقة تشكيلها للجمل وتعبيرها عن الأفكار، فما يعبَّر عنه في جملة واحدة باللغة الإنجليزية قد يعبَّر عنه في جملتين في اللغة العربية، وهنا تواجه برامج الترجمة الآلية الإحصائية مشكل حدوث الخطء عند محاذاة النصوص المدخلة وخاصة في حال محاذاة نصين من أزواج لغوية متباعدة (O Oransa.، 2019) ذلك لأن عملية المحاذاة تتم بطريقة آلية ولا تخضع غالبا للمراجعة والتدقيق البشري.
- ترجمة التعابير الاصطلاحية (idioms) : تعجز أداوت الترجمة الآلية الإحصائية عن ترجمة التعابير الاصطلاحية بشكل سليم يتوافق وثقافة اللغة الهدف ذلك لأنها تتشكل من مجموعة من الكلمات التي لا يتضح معناها إلا بتجميعها في إطار محدد، فمثلا : يشيع في البرلمان الكندي استعمال عبارة « hear ! hear ! » للتعبير عن إعجاب الحاضرين بكلام المتحدث. حيث بدأ استعمال هذا التعبير الاصطلاحي في القرن السابع عشر كوسيلة لإبداء موافقة أعضاء مجلس العموم (House of Commons) على رأي المتحدث بدلا من التصفيق الذي كان ممنوعا وقتها في الاجتماعات الرسمية. وبدلا من أن يترجم برنامج الترجمة الآلية الإحصائية العبارة إلى « أحسنت ! أحسنت ! » يترجمها حرفيا إلى « اسمع ! اسمع ! ».
- ترتيب الكلمات (order word) : تختلف طريقة ترتيب كلمات الجملة من لغة إلى أخرى وبالأخص إذا كانت قواعد اللغة المنقول منها وإليها تختلفان بشكل كبير حال اللغة العربية والإنجليزية، مما قد يؤدي بأداة الترجمة الآلية الإحصائية إلى ترجمة الجمل ترجمة خاطئة من حيث ترتيب الكلمات في الجملة الواحدة كتقديم الاسم على الفعل في تشكيل الجمل الفعلية في اللغة العربية. ويزداد الأمر صعوبة كلما زاد طول الجملة وتداخل أفكارها.

- الأخطاء النحوية (syntactic errors) : عند ترجمة النصوص على أداة الترجمة الآلية الإحصائية يمكن ظهور مشاكل نحوية متعلقة بسوء تصريف الأفعال أو عدم التمييز بين المثنى والجمع أو بين التأنيث والتذكير، ويرجع سبب ذلك إلى عدم تدريب المتون أحادية اللغة بشكل دقيق ومدروس أو عدم إضافة معلومات نحوية كافية للذخيرة اللغوية المخزنة في قاعدة بيانات البرنامج.

5. الترجمة الآلية العصبية : الماهية والأهمية

سلكت الترجمة الآلية في الآونة الأخيرة نهجا جديدا باستبدالها للمنهج الإحصائي بالنموذج العصبي واعتمادها على تكنولوجيا الذكاء الاصطناعي لترجمة النصوص. وتعتمد الترجمة الآلية العصبية على طريقة عمل جديدة تُعرف بنموذج التشفيروفك التشفير (Model Decoder-Encoder) (Kyunghyun 2014 :1) من خلال بناء وتدريب شبكة عصبية واحدة كبيرة تحلل الجمل المدخلة كاملة وتترجمها إلى اللغة الهدف بشكل موحد (Bahdanau et al. 2015 :1)، وبعبارة أخرى يتشكل برنامج الترجمة الآلية العصبية النموذجي من شبكتين عصبيتين أساسيتين، أولهما تعالج جمل النص المدخل والأخرى تولد الجمل المترجمة عن طريق النواقل (vectors) (Bahdanauوal. 2015 :1) . ولهذا يسمى منهج الترجمة الآلية العصبية بمنهج الطرفين (Approach End-to-End).

وتتكون الشبكة العصبية الاصطناعية من وحدات معالجة مكونة من عناصر حسابية تسمى عصبونات (neurons) مهمّتها تخزين المعلومات ومعالجتها بطريقة شاملة ومتصلة. ويشبّه الباحثون الشبكة العصبية الاصطناعية بدماغ الطفل الذي يُولد دون أدنى فكرة عما يدور حوله ولكنه يبدأ باكتساب المعارف وتخزينها ثم استرجاعها عند الحاجة، غير أن ما يميز الشبكة العصبية الاصطناعية عن الشبكة العصبية البشرية -وإن كانت أقل تعقيدا -أنها لا تنسى ما اكتسبته، فكل معلومة تتلقاها الشبكة العصبية الاصطناعية تخزَّن وتُحفظ إلى الأبد (Krzysztof 2015 :2) كما أن محاكاة العقل البشري في تصميم نظم الترجمة الآلية العصبية يساعد على تعليم خوارزميات البرنامج كيفية التعلم من الخطأ والتكرار، وكيفية التذكروالقدرة على اتخاذ القرارالسليم وكذا تخزين المعلومات على شكل ملفات منظمة واسترجاعها وقت الحاجة، فإذا قام المستخدم بتعديل الذخيرة اللغوية أو تصحيح خطأ ما فإن البرنامج يتأقلم مع هذا التغيربطريقة سريعة وتلقائية. وعليه، يتم تقسيم الشبكة العصبية الاصطناعية إلى طبقات ثلاث : طبقة المدخلات (inner layer)، طبقة المخرجات (outer layer) والطبقة الخفية (hidden layer)، ويتصل كل عصبون بما يسبقه وما يليه من العصبونات في الطبقات الأخرى .(Koehn 2015 : 9)



شكل (02) : طبقات الشبكة العصبية الاصطناعية الثلاث

وتتفوق نظم الترجمة الآلية العصبية على نظيرتها الإحصائية من حيث سعة تخزين المعطيات (Kyunghyun 2014 :1)، كما أنها تُعدّ أكثر مناهج الترجمة كفاءة إلى حد الآن كونها تركز على سياق النص كاملا وتتعامل معه على أنه كتلة واحدة، وهذا ما قد يختصر طريق المترجم ويقلص من الوقت المخصص للمراجعة والتعديل، حيث تقدَّر نسبة المراجعة في نظم الترجمة الآلية العصبية بأقل من 26 % (Bentivogli 2016 : 265) مقارنة ببرامج الترجمة الآلية الإحصائية التي تصل نسبة المراجعة فيها إلى 30 % (Koehn 2010 : 231).

6. عوامل نجاح الترجمة الآلية العصبية

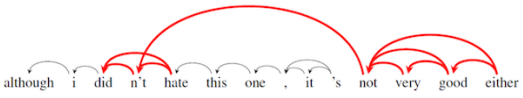
إن أكثر ما يميز الترجمة الآلية العصبية عن الترجمة الآلية الإحصائية اعتمادها على جملة من الآليات التكنولوجية التي حسنت من أداء خوارزمياتها على مستوى التخزين وأمن المعلومات والسرعة في الأداء، أبرزها :

- التعلم العميق (deep learning) : بدأ استعمال آلية التعلم العميق في برامج الترجمة الآلية العصبية في ضل توفر المزيد من البيانات الرقمية التي يمكن استثمارها في تدريب الشبكات العصبية الاصطناعية. ويتم تصميم نماذج التعلم العميق وتدريبها بتمثيل بيانات التعلم بدلا من خوارزميات محددة وإضافة

العديد من الطبقات الخفية بين طبقة المدخلات وطبقة المخرجات، وكلما زاد عدد الطبقات الخفية كلما زاد التعلم عمقا وتعقيدا (Chaudhary et al : 2018 : 145). وفي بحث قامت به (Hardmeier و Vanmassenhove 2018 : 3007) لدراسة كيفية تعامل الترجمة الآلية العصبية مع اختلاف جنس التذكير والتأنيث بين اللغة الإنجليزية والفرنسية، توصلتا إلى أن تحديد جنس الأشخاص في الجملة إما عن طريق ذكر أسماء العلم (John, Marry)، أو بإضافة الضمائر (he, she) أو النعوت (heureuse, envahissante) يساعد على تحسين جودة ترجمة النص كاملا وليس الجمل الخاضعة للتعديل فقط، وهذا ما يؤكد فعالية مبدأ التعلم العميق في الترجمة الآلية.

- التوقع (prediction) : يتم تصميم نظم الترجمة الآلية العصبية بطريقة تمكنها من تطوير آلية التوقع واستغلالها في مرحلة الترجمة، حيث يتم تدريب البرنامج بنفس الطريقة التي يتم فيها تدريب أجهزة إكمال النصوص (text completion devices) بطريقة تلقائية وسريعة كما يظهر في لوحة مفاتيح الهواتف الذكية (Forcada 2017 : 295)، فعندما يكتب المستخدم كلمة ما على الهاتف يقترح عليه الجهاز الكلمة الموالية بناءً على معنى الكلمة السابقة. وتعمل نظم الترجمة الآلية العصبية بنفس الطريقة ولكن على نحو أكثر تعقيدا، فعند إدخال النص للترجمة يبدأ البرنامج بالبحث عن جميع الكلمات المرشحة لترجمة الكلمة الأولى، ثم ينتقل إلى الكلمة الثانية انطلاقا من معنى الكلمة السابقة (Zong 2018 : 5) وهكذا دواليك إلى أن ينهي ترجمة محتوى النص كاملا.
- الانتباه (attention) : كانت نظم الترجمة الآلية العصبية في بداية ظهورها تعجز تماما عن ترجمة الجمل الطويلة ومعقدة التركيب بسبب فشل النواقل (vectors) في تعيين التسلسل الصحيح لعناصر الجملة في اللغة الهدف، وهذا ما دفع الباحثين لتطوير آلية الانتباه لتمكين البرنامج من ترجمة الجمل الطويلة بسهولة وتركيب عناصرها حسب قواعد وخصوصيات اللغة الهدف.

مثال :



يتضح في هذا المثال كيفية عمل آلية الانتباه، حيث يتم ربط كل عنصر من عناصر الجملة الأصل (حتى علامات الوقف) بما يسبقها لتحديد وظيفتها وسياقها. وتجرى عملية الانتباه في هذا المثال على مستويين، يربط المستوى الأول الكلمات بما يسبقها مباشرة في الجملة، مثل : « although مع I »، على عكس المستوى الثاني الذي يربط الكلمات المتشابهة من حيث الوظيفة حتى ولو كانت متباعدة في الجملة، مثل أدوات النفي : « not... either » مع « didn't ».

غير أن النموذج العصبي في الترجمة الآلية وعلى الرغم من شهرته وانتشار استعماله في الوقت الراهن، لازال يطرح بعض المشاكل اللغوية والتقنية الجوهرية، أهمها صعوبة وبطأ عملية تدريب النماذج كونها تستدعي تغذية البرنامج ببيانات ضخمة ثم معالجتها ومحاذاة المتون الموازية في اللغة الأصل والهدف، ويحتاج كل ذلك للكثير من الوقت والجهد وتوفير مزيد من الإمكانيات المادية والتقنية والبشرية المهيأة لتنفيذ هذه المهام المعقدة والدقيقة. من جهة أخرى تفشل نظم الترجمة الآلية العصبية في التعامل مع الكلمات النادرة والمصطلحات الجديدة وأسماء العلم، فتترجمها ترجمة خاطئة أحيانا أو لا تترجمها أصلا (Yonghui et al. 2016 : 32)، بالإضافة إلى عجزها عن الربط بين جمل وفقرات النص باستعمال أدوات الربط المناسبة أو الجمع بين الجمل أو تقسيمها بما يناسب طريقة التعبير في اللغة الوصل.

7.الدراسة المقارنة

انطلاقا من دراستنا النظرية لموضوع الترجمة الآلية الإحصائية والعصبية، ارتأينا ترجمة نص متخصص من اللغة الإنجليزية إلى العربية -يضم عددا من المصطلحات والاختصارات والمفاهيم العلمية-على أداةReverso القائمة على المنهج الإحصائي وأداة Google Translate القائمة على النموذج العصبي، ثم مقارنة نِتاجهما من خلال رصد الأخطاء النحوية والتركيبية والمعجمية والدلالية في كلتا النسختين.

1.7.النص

What's new in machine translation ? Today, it is no exaggeration to say that technological developments in translation, driven by the two major technological innovations of computer-assisted translation tools and machine translation, have fundamentally changed how we communicate. While the introduction of AI to translation may still be in its infancy, it has already achieved promising results ; NMT has just landed,

and with the involvement of deep learning, it is here to stay. Real-time translation has also experienced a quantum leap in the world of business and audiovisual industry and tends to move us closer to 'language transparent'society. Besides, cloud machine translation enables translators to collaborate on projects in real time and share resources in the easiest way. However, we are still not sure whether, in near future, neural machine translation can hold a candle to the depth of experience and bicultural understanding that a human translator applies to translating the source text**.**

2.7.نتاج الترجمة الآلية

1.2.7. المستوى النحوي والتركيبي

technological that say to exaggeration no is It ,Today innovations	
✖ Reverso (SMT) : اليوم، ليس من المبالغة أن نقول بجرأة إن التطورات التكنولوجية ✖ Translate Google (NMT) : اليوم، ليس من المبالغة أن نقول بجرأة أنّ التطورات التكنولوجية	
communicate we how changed fundamentally have	
✔ Reverso (SMT): غيرت بشكل جوهري طريقة تواصلنا ✖ Translate Google (NMT): غيرت بشكل أساسي كيف نتواصل	
the to candle a hold can translation machine Neural depth...	
✖ Reverso (SMT): أن الترجمة الآلية من الممكن أن تحمل ✔ Translate Google (NMT) يمكن أن تحمل الشمعة على عمق :	
text source the translating when	
✖ Reverso (SMT): عندما يترجم النص المصدر ✔ Translate Google (NMT): عند ترجمة النص المصدر	
technological major two the by driven	
✖ Reverso (SMT): مدفوعة بالابتكارات ✔ Translate Google (NMT): مدفوعة بالابتكارات	

in be still may translation to AI of introduction the While ...infancy its	
✖ Reverso (SMT) : ورغم أن إدخال الذكاء الاصطناعي إلى الترجمة ربما لا يزال ✔ Translate Google (NMT) : في حين أن إدخال الذكاء الاصطناعي للترجمة قد لا يزال	
results promising achieved already has it	
✖ Reverso (SMT) : فإنه حقق بالفعل نتائج واعدة ✔ Translate Google (NMT) : فقد حقق بالفعل نتائج واعدة	
way easiest the in resources share	
✖ Reverso (SMT) : مشاركة الموارد بالطريقة الأسهل ✔ Translate Google (NMT) : مشاركة الموارد بأسهل طريقة	

2.2.7. المستوى المعجمي

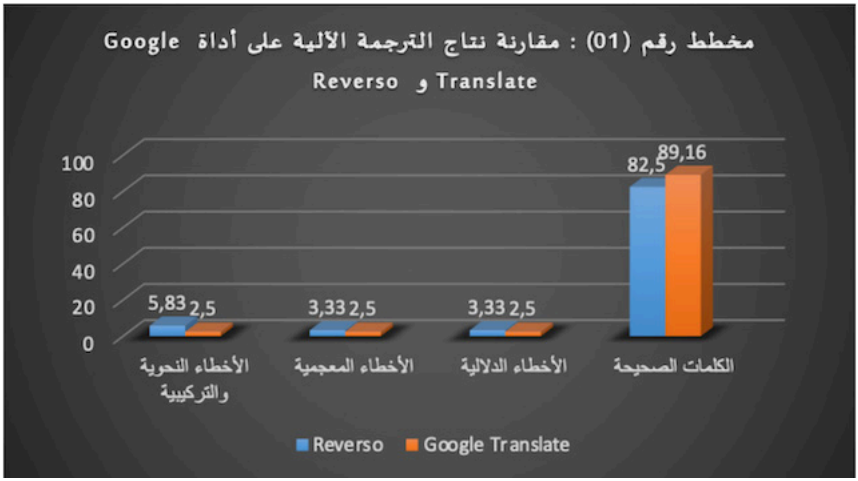
have fundamentally changed how we communicate.	
✔ Reverso (SMT) : غيرت بشكل جوهري طريقة تواصلنا ✖ Google Translate (NMT) : غيرت بشكل رئيسي كيف نتواصل	
While the introduction of AI to translation…	
✖ Reverso (SMT) : إدخال الذكاء الاصطناعي ✖ Google Translate (NMT) : إدخال الذكاء الاصطناعي	
Real-time translation has also experienced a quantum leap…	
✔ Reverso (SMT) : شهدت قفزة نوعية ✔ Google Translate (NMT) : شهدت قفزة نوعية	
tends to move us closer…	
✖ Reverso (SMT) : تميل إلى تقريبنا ✖ Translate Google (NMT) : تميل إلى تقريبنا	
collaborate on projects in real time	
✖ Reverso (SMT) : التعاون في المشاريع في الوقت الحقيقي ✔ Translate Google (NMT) : التعاون في المشاريع في الوقت الفعلي	
the depth of experience	
✖ Reverso (SMT) : عمق الخبرة ✔ Google Translate (NMT) : عمق التجربة	

3.2.7. المستوى الدلالي

translation assisted -computer	
Reverso (SMT) : الترجمة بساعدة الحاسوب	✔
Translate Google (NMT) : الترجمة بمساعدة الكمبيوتر	✖
AI	
Reverso (SMT) : الذكاء الاصطناعي	✔
Translate Google (NMT) : الذكاء الاصطناعي	✔
its in be still may translation to AI of introduction the While infancy	
Reverso (SMT) : ربما لا يزال في مرحلة الطفولة المبكرة	✖
Translate Google (NMT) : قد لا يزال في بدايته	✔
landed just has NMT	
Reverso (SMT) : فقد هبط فريق التدريب الوطني للتو	✖
Translate Google (NMT) : للتو NMT لقد هبطت	✖
experienced also has translation time-Real...	
Reverso (SMT) : شهدت الترجمة في الوقت الحقيقي	✖
Translate Google (NMT) : شهدت الترجمة الفورية أيضًا	✔
candle a hold can translation machine neural	
Reverso (SMT) : من الممكن أن تحمل شمعة	✖
Translate Google (NMT) : يمكن أن تحمل شمعة	✖

جدول (02) نتاج ترجمة Reverso وGoogle Translate بالنسب المئوية

Reverso				
الأخطاء	المستوى النحوي– التركيبي	المستوى المعجمي	المستوى الدلالي	الكلمات الصحيحة
العدد	7	4	4	99
النسبة المئوية	5,83 %	3,33 %	3,33 %	82,5 %
Google Translate				
الأخطاء	المستوى النحوي-التركيبي	المستوى المعجمي	المستوى الدلالي	الكلمات الصحيحة
العدد	3	3	3	107
النسبة المئوية	2,5 %	2,5 %	2,5 %	89,16 %



4.2.7. التحليل

بعد دراسة جودة ترجمتَي Reverso وGoogle Translate ومقارنتهما من حيث عدد الأخطاء النحوية والتركيبية والمعجمية والدلالية ونسبة الكلمات الصحيحة، تبيّن لنا أن Google Translate تفوقت على Reverso في جميع الجوانب ولكن بنسب قليلة وذلك لحداثة نموذج الترجمة الآلية العصبية مقارنة بمنهج الترجمة الآلية الإحصائية. ويتّضح لنا من خلال تحليل النص المترجم على Reverso (المنهج الإحصائي) أنّ الأداة لا تعكس مبدأ الترجمة الآلية بالضبط، ذلك لأن ما تقوم به هو البحث في قاعدة بياناتها عن جمل مشابهة لما يُعرض عليها وتستخدم المخزون كترجمة للنُّصوص المعروضة، على عكس النص المترجم على Google Translate (النموذج العصبي) الذي نلحظ فيه بعض الاجتهاد والإبداع في الترجمة وذلك لأن البرنامج يعتمد في الأساس على تكنولوجيا التعلم العميق وتقنية الانتباه والتوقع...إلخ.

- من الناحية النحوية: تتعلق المشاكل النحوية الصادرة عن Google Translate و Reverso بترتيب الكلمات وتحديد أزمنة الأفعال واختيار أدوات الربط السليمة، حيث أن Google Translate كانت أكثر دقة في ترجمة الفعل المضارع (neural machine translation can hold ← يمكن أن تحمل) والمضارع التام (has it already achieved promising results ← فقد حقق بالفعل نتائج واعدة). إلى جانب ذلك، نجحت Google Translate في ترجمة أداوت الربط وحروف الجر ترجمة سليمة، وهذا بفضل آلية التعلم العميق الذي تخضع له الشبكة

العصبية الاصطناعية مما يساعدها على اكتساب معلومات جديدة وتكييف بياناتها بما يناسب خصوصيات وتراكيب اللغة الهدف، على عكس Reverso التي فشلت في ترجمة 'of'بحرف الجر 'في' بدلا من 'ل' في الجملة (the two major technological innovations of computer-assisted translation tools and machine translation ← الابتكارات التكنولوجية الرئيسية في أدوات الترجمة بمساعدة الحاسوب والترجمة الآلية)، وكذا أداة الربط 'that'التي ترجمتها إلى 'إن' بدلا من 'أنّ' في الجملة← (it is no exaggeration to say that technological ليس من المبالغة أن نقول إن التطورات). علاوة على ذلك فشلت كلتا الأداتين في ترتيب بعض جمل النص ترتيبا يتوافق مع قواعد بناء الجملة الفعلية في اللغة العربية (←today, it is no exaggeration اليوم، ليس من المبالغة) كما فشلتا في مطابقة النعت والمنعوت في (technological developments in translation, driven by ← التطورات التكنولوجية في الترجمة، مدفوعة).

- من الناحية المعجمية : نجحت Google Translate في ترجمة الكثير من كلمات النص ترجمة سليمة مناسبة لسياق النص (time real / experience التجربة/ الوقت الفعلي) غير أنها فشلت في ترجمة الكلمات التي تحمل معنيين مختلفين كالفعل 'to tends' الذي يعني 'التحرك نحو اتجاه معين' أو 'احتمال امتلاك خاصية معينة'. حيث اختارت Google Translate المعنى الأول الذي لا يتوافق مع سياق النص (tends to move us closer to language transparent society تميل إلى تقريبنا من مجتمع اللغة الشفافة)، ذلك لأن أتمتة (automation) مرحلة فك التشفير في نظم الترجمة الآلية العصبية قد تضر بالنص المترجم أكثر من أن تخدمه. ولقد لاحظنا نفس هذا التباين في ترجمة Reverso القائمة على الاحتمالات، فكلما نقص تنوع وحجم المتون الموزاية المخزنة في قاعدة بياناتها كلما تضاءلت نسبة نجاح ترجمة الكلمات ترجمة سياقية مقبولة. ارتكبت Google Translate و Reverso الخطأ نفسه في ترجمة الكلمة 'involvement'التي تُرجمت من قِبل الأداتين إلى 'إشراك' (← with the involvement of deep learning وبإشراك التعلم العميق)، والأصحّ ترجمة الكلمة إلى 'استعمال أو استثمار' للتعبير عن فكرة استخدام آلية التعلم العميق في نظم الترجمة الآلية العصبية.
- من الناحية الدلالية : فيما نجحت Google Translate في ترجمة تعبير مجازي واحد بشكل صحيح من أصل ثلاثة (← in its infancy لايزال في بدايته)، ترجمت Reverso جميع التعابير الاصطلاحية والاستعارات الواردة في النص ترجمة حرفية خاطئة وذلك بسبب اللبس اللغوي الذي يكتنف هذه العبارات

وعجز الخوارزميات عن تحديد معناها الخفي وإيجاد مكافئها الثقافي في اللغة الوصل بطريقة آلية. فمن وحي دراستنا المقارنة، توصلنا إلى أن أدوات الترجمة الآلية الإحصائية والعصبية لازالت عاجزة عن استيعاب معنى التعابير المجازية على عكس المترجم المحترف الذي يتجاوز المعنى الظاهر للكلمة ويتعمق في معانيها الخفية مستعينا بسياق النص للوقوف على كُنه العبارة وما يتفرع عنها من ضلال في المعنى. إلى جانب ذلك، وُفّقت Google Translate في ترجمة معظم مصطلحات النص (الذكاء الاصطناعي، الترجمة الفورية) غير أنها في المقابل لم تنجح في التعرف على اختصار 'NMT' وأعادت نقله إلى النص الهدف كما هو، في حين ترجمته Reverso ترجمة بعيدة عن مجال تخصص النص 'فريق التدريب الوطني'. ويعود سبب ذلك على الأرجح إلى أن مفهوم الترجمة الآلية العصبية لازال حديث الاستعمال على شبكة الأنترنت، مما يبطئ عملية إدخاله في قاعدة بيانات نظم الترجمة الآلية الإحصائية والعصبية.

الخاتمة

مما لا ريب فيه أن التكنولوجيات الحديثة غيرت مسار الترجمة تنظيرا وممارسة وأدخلت عليها مفاهيم وأدوات لم يكن للمجتمعات السابقة علم بها. وما لا يسع المترجمَ المحترف جهلهُ تنوع خدمات تكنولوجيات الترجمة وتشعب مهامها وبالأخص نظم الترجمة الآلية التي أبانت عن جدارتها في توفير خدمات قيمة وحلول عملية يصعب على المترجم اليوم تجاهلها أو إنكار فعاليتها. ولقد حاولنا في هذه الدراسة التركيز على دوافع استبدال منهج الترجمة الآلية الإحصائية بنموذج الترجمة الآلية العصبية ودراسة هذه الإشكالية من الناحية النظرية والتطبيقية وما لهذا التغير من تأثير على جودة الترجمة المتخصصة.

ولقد أفضت هذه الدراسة إلى أن المجتمع البحثي في العالم الغربي والعربي يعلق آمالا عِراضا على أدوات الترجمة الآلية العصبية ويرى فيها ثمرة الأبحاث والتجارب التي شهدتها نظم الترجمة الآلية منذ أربعينيات القرن الماضي إلى يومنا هذا. غير أن هذا النجاح لا يكتمل إلا بتطوير الجانب الأكاديمي لتكنولوجيات الترجمة من خلال طرح المسائل الشائكة والمستجدة في الملتقيات وتدريب طلبة الترجمة على استثمار خدمات الترجمة الآلية على أكمل وجه. كما أن تحسين أداء الترجمة الآلية إلى اللغة العربية ونجاح العالم العربي في معالجة اللغة العربية بطريقة آلية مرهون بتعزيز المحتوى الرقمي المنشور بها من جهة وتصميم معجم عربي محوسب وموحد يجمع كلمات اللغة العربية ومقابلاتها في اللغات الأجنبية من جهة أخرى ليسهل التعامل مع مفرداتها في كافة التطبيقات ذات العلاقة

بالترجمة الآلية. وصفوة القول أنّ مشاريع الترجمة الآلية اليوم أصبحت أكثر جدّية وكفاءة مما كانت عليه من قبل بفضل ما توصل إليه العلم من تقانة وحداثة غير أنها وبوضعها الحالي لازالت قاصرة على تقديم تراجم مقبولة شكلا ومضمونا إلا بتدخل المترجم البشري في مرحلة التحرير السابق واللاحق لتصبح قابلة للنشر والقراءة.

باللغة العربية

بن تركي، وفاء، سمار، نصر الدين. 3102. « اختبار أداء نظام الترجمة الآلية الإحصائية المكيّف لدعم الثنائية اللغوية إنجليزي-عربي» Moses . Revue de l'Information Scientifique et Technique. العدد 2. المجلد 02. ص 611-531.

حجازي، محمود. 5991. الأسس اللغوية لعلم المصطلح، ط1، القاهرة، دار غريب.

كروجر، رالف، ترجمة فؤاد، لبنى، 8102، تعدد اللغات في ضوء الترجمة الآلية https://www.goethe.de/ins/eg/ar/spr/mag/21250571.html (consulté le 8 février 2020)

قائمة المراجع

باللغة الأجنبية

Alrajeh ,A .2018 .*A Recipe for Arabic-English Neural Machine Translation*. Riyadh : National Center for AI and Big Data Technology.

Bahdanau ,D ,.Kyunghyun ,C ,.et al" .2015 .Neural machine translation by jointly learning to align and translate ."Proceedings of the international conference on learning Representations .San Diego .USA.

noitalsnart enihcam desab-esarhp susrev larueN" .6102 .L ,ilgovitneB no ecnerefnoC lanoitanretnI eht fo sgnideecorP ,"yduts esac a : ytilauq .ASU .saxeT .gnissecorP egaugnaL larutaN ni sdohteM laciripmE

Castaño ,A ,.Casacuberta ,F .et al" .1997 .Machine translation using neural networks and finite-state models ."Proceedings of the7 th International Conference on Theoretical and Methodological Issues on Machine Translation. New Mexico .USA .pp.167- 160

Chaudhary ,J ,.Ankit ,C" .2018 .Machine translation using deep learning: a survey ."IJSRSET Journal .(02) 04 .pp.150 – 145

Forcada ,M" .2017 .Making sense of neural machine translation." Translation Spaces .(2) 6 .pp.309 – 291

Jelinek F ,.Brown ,P ,.et al" .1990 .A statistical approach to machine translation ."The Journal of Computational Linguistics .(2) 16 pp.85 -79

Koehn ,P .2010 .*Statistical Machine Translation* .Cambridge : Cambridge University Press.

Koehn ,P .2015 .*Neural Machine Translation* ,London : Center for Speech and Language Processing.

Krzysztof ,W ,.Krzysztof ,M" .2015 .Neural-based machine translation for medical text domain ,"Procedia Computer Science Journal .(64) pp.2-9

Kyunghyun C" .2014 .On the properties of neural machine translation: encoder – decoder approaches ."Proceedings of the8 th Workshop on Syntax. Semantics and Structure in Statistical Translation .Doha .Qatar.

Oransa ,O .2019 .statistical machine translation

http://osama-oransa.blogspot.com/2012/01/statistical-machine-translation-smt.html) Consulté le16 février 2020.(

Vanmassenhove ,E ,.Hardmeier ,C .et al" .2018 .Getting gender right in neural machine translation ."Proceedings of the International Conference on Empirical Methods in NLP .Brussels .Belguim .pp.3003-3008

Yonghui ,M ,.Zhifeng ,Ch .et al .2016 .Google's neural machine translation system : bridging the gap between human and machine translation

https://arxiv.org/abs/1609.08144) consulté le20 février 2020.(

Yu L ,.Zong C" .2018 .Exploiting knowledge graphs in neural machine translation ."Proceedings of the14 th China Workshop on Machine Translation .Wuyishan .China.

مستخلص

إن المتطلع على مسار تطور أداوت الترجمة الآلية يجد أنها شهدت تغيرات عدة تُوّجت باعتماد آلية الشبكات العصبية الاصطناعية في تطوير تصميماتها وتبني تقنيات التعلم العميق في تحسين أداء خوارزمياتها. وكغيرها من لغات العالم، لم تكن اللغة العربية في معزل عن هذا التغير، بل استفادت من هذه الأدوات التكنولوجية في إنعاش ميدان الترجمة العامة والمتخصصة.ومنه، حاولنا في دراستنا المقارنة هذه تحري طبيعة أداوت الترجمة الآلية العصبية وتلك المعتمدة على المناهج الإحصائية، ومدى تأثير كل منهما على جودة النص المتخصص المترجم من اللغة الإنجليزية إلى العربية، للخروج بجملة من

النتائج مفادها أن الترجمة الآلية العصبية قدمت نتائج واعدة من حيث سلامة اللغة المتخصصة ودقة جهازها المصطلحي، ولكنها لم تصل بعد إلى مستوى الترجمة البشرية.

كلمات مفتاحيّة

الترجمة الآلية؛ النص المتخصص؛ الجودة؛ اللغة الإنجليزية؛ اللغة العربية

Résumé

La traduction automatique fasse à notre époque l'objet d'un intérêt grandissant, en particulier après l'utilisation de réseaux de neurones artificiels la performance de la et les mécanismes d'apprentissage profond pour améliorer traduction automatique neuronale. La langue arabe - comme les autres langues du monde - a profité de ce changement pour améliorer la traduction générale et spécialisée.

Notre étude comparative vise à étudier les motifs et l'impact de ce changement sur la qualité de la traduction automatique de l'Anglais vers l'Arabe ; pour conclure que bien que la traduction automatique neuronale ait donné des résultats prometteurs au niveau linguistique et terminologique, elle n'est pas comparable à la traduction humaine.

Mots-clés

Traduction Automatique ; Texte Spécialisé ; Qualité ; Anglais ; Arabe

Abstract

Machine translation is of increasing interest in our time, especially after the use of artificial neural networks and the adoption of deep learning mechanisms in of neural machine translation performance. Arabic language – the improvement like other world languages – benefited from this change to improve the process of general and specialized translation.

Our comparative study explores the motifs and the impact of this change on the quality of English to Arabic machine translation ; to conclude that although neural machine translation has produced promising linguistic and terminological results, the raw output is not comparable to human translation.

Keywords

Machine Translation ; Specialized Text ; Quality ; English ; Arabic