استثمار المدونات المتوازية في تطوير حوسبة اللغة العربية

Using Parallel Corpora in Developing Arabic Computing

أسامة شتيوي[1]، فتيحة داني[2]

Oussama CHETIOUI[1], Fatiha DANI[2]

[1] جامعة وهران 1، (الجزائر)، oussama.chetioui@edu.univ-oran1.dz

[2] جامعة وهران 1، (الجزائر)، fatiha.dani@yahoo.com

تاريخ الاستلام: 2021/07/05 **تاريخ القبول**: 2021/11/01 **تاريخ النشر**: 2021/12/28

ملخص:

لم يعد مفهوم حوسبة اللغة العربية مقتصرا على الحرف العربي وسبل برمجته في الحاسوب، فقد قطعت اللغة العربية أشواطا كبرى في هذا الإطار. لكن المسألة حاضرا تتعلق بجعل الحاسوب أكثر تناغما مع الدلالات العربية في السياقات المختلفة، ومساعدته على توليدها في شتى السياقات بما يتناسب مع نُظم اللغة العربية. واقترحت هذه الدراسة نهجا سياقيا متمحورا حول استثمار ما تقدمه المدونات المتوازية من سياقات لغوية وتُرجمية في سبيل الدفع بهذا المسعى قدما. وحاولت الدراسة عمليا تبيان مدى فاعلية الاعتماد على ذلك التوازي المقارن بين سياقي النصين في المدونة المتوازية الواحدة من أجل تعريف برمجيات الحاسوب أكثر بسبل معالجة اللغة العربية للسياقات العامة والمتخصصة. وقد تبين من خلال طرح عدد من الأمثلة التطبيقية الفروق البينة بين البرمجيات المعتمدة على هذا النهج السياقي وأخرى لا تعتمده. وخلصت الدراسة إلى تقديم حجج داعمة لهذا النهج السياقي في سبيل تطوير حوسبة اللغة العربية.

كلمات مفتاحية: الحوسبة العربية، اللسانيات الحاسوبية، النهج السياقي، المدونات المتوازية، الدلالات العربية.

Abstract:

The concept of Arabic computing is no longer based on its alphabet programming; Arabic language has already advanced in this context. The issue, nowadays, is on familiarizing the computer with Arabic denotations in various contexts, attempting to generate them in the appropriate language structure. Doing so, this study proposes a contextual method benefitting of parallel corpora provided by lingual and translational contexts. This study has, practically, examined the effectiveness of this comparable parallelism between both texts' contexts to familiarize the computer software with the Arabic processing of general and specialized contexts. After examining some practical examples, clear-cut differences emerged between applying and non-applying translation software of this contextual method. This study concludes by supporting the contextual method as an adoptable mechanism in developing Arabic computing.

Keywords: Arabic Computing; Computational Linguistics; Contextual Method; Parallel Corpora; Arabic Denotations.

المؤلف المرسِل: أسامة شتيوي، **الإيميل**: oussama.chetioui@edu.univ-oran1.dz

1. مقدمة:

إن ما تناقشه هذه الدراسة يتجاوز مرحلة ادخال اللغة العربية في آليّات الحاسوب. فاللغة العربية قد أحرزت بالفعل تقدما معتبرا بهذا الصدد. فأصبحنا نرى الخط العربي بأشكاله المتعددة بشكل اعتيادي في برامج الحاسوب، وبات الحاسوب اليوم قادرا على التعرف على الحركات الاعرابية ومواضع الهمزة وسبل ورودها في الكلمة، بل وبات يصحح الخطأ آليّا أو يشير إليه على أقل تقدير. وهي خصوصيات طالما شكلت صعوبة في حوسبة اللغة العربية بادئ الأمر. فالدراسة إذا تناقش سبل تطوير تعرف الحاسوب على الدلالات العربية في سياقاتها الأصلية عبر ما يسمى بالسياق الترجمي، أي أن يتمكن الحاسوب من معالجة الدلالات العربية المدرجة في برمجياته وتحليل شحناتها العامة والمتخصصة على حد سواء. والترجمة بوصفها حقلا سياقيا بامتياز تؤدي دورها في توفير سياقات مختلفة الانتماءات بين ما هو عام ومتخصص، في النصوص المترجمة إما آليّا أو بشريا، والمدرجة في برامج الترجمة الآليّة كمدونات داعمة لذاكرة الترجمة. في هذا السياق، تركز هذه الدراسة على المدونات المتوازية بصفتها أداة ناجعة فيما تقدمه من سياقات لغوية متوازية، من شأنها أن تسهل على الحاسوب ربط السياقات بعضها ببعض عبر فعل التوازي في السطر البرمجي الواحد. فعند ادراج مدونة متوازية انكليزية – عربية مثلا في برنامج أو خدمة للترجمة الآليّة، فإننا بذلك نخلق سياقا ترجميا جديدا يُحفظ في ذاكرة الترجمة، التي تعتمدها حاليا البرامج والخدمات الترجمية الآليّة بأنواعها كافة. ما يحدث هنا أن يدخل ذلك السياق الجديد، والمتولد من اعتماد المدونة المتوازية في ترجمة النص قيد الخدمة، في برمجيات الذكاء الاصطناعي لذلك البرنامج، والذي يحفظه بدوره تلقائيا ليُستخدم فيما بعد في سياق ترجمي مماثل. فهو بذلك يستخرج المصطلحات المتطابقة سياقيا ويدرجها في السياق الجديد مباشرة، ما من شأنه أن يرفع من مستوى الجودة في النص المتولد ويخدم خصوصية السياق السائد في العملية برمتها.

وفي إطار السعي الحثيث نحو الوصول إلى حوسبة عربية أكثر تطورا، برزت جهود عدة تبنت هذا السعي وحاولت تعريف الحاسوب أكثر بنُظم اللغة العربية وبخصائصها اللغوية والتركيبية وما إلى ذلك من الأبعاد البنيوية للغة العربية. وهو مسعاً لطالما شغل اهتمام الباحثين في اللسانيات الحاسوبية العربية على وجه التحديد. إلا أنه وعلى الرغم من كل تلك الجهود، لا تزال اللغة العربية تعاني قصورا في حوسبتها مقارنة بما وصلت إليه لغات جمة من مستويات. وهو ما يشير إلى احتمالية قصور المنهج المتبع في العملية برمتها. فاللغة العربية لها ما يميزها الكثير عن غيرها من اللغات، وبدا منطقيا اعتماد منهج مختلف في سبيل حوسبتها عن ذلك الذي اعتمدته اللغات الأخرى ذات السبق في المجال. فهي لغات تختلف اختلافا جذريا في بنيتها عن اللغة العربية، ما يعني ضرورة اعتماد منهج فريد يستوعب اللغة العربية بكل مكنوناتها، ويُسهم في الدفع بحوسبتها قدما نحو مواكبة التطور الحاصل في اللسانيات الحاسوبية وما باتت تعكف عليه من اسهامات تقدمها في شتى المجالات. وعليه، تنطلق هذه الدراسة من التساؤل المحوري الآتي: ما هو واقع حوسبة اللغة العربية مقارنة بباقي اللغات؟ وفي محاولة للإجابة على هذا الاشكال، تطرح الدراسة ما يلي من تساؤلات فرعية: كيف يؤثر تأخر حوسبة اللغة

العربية سلبا على مواكبتها للتطور الحاصل في المجال؟ وهل للمدونات المتوازية ما تقدمه من نُهج في سبيل الدفع بالحوسبة العربية قدما؟

وسعيا للإجابة على هذه التساؤلات، تناقش هذه الدراسة واقع حوسبة اللغة العربية وسُبل تطويره، عبر الاستثمار في المدونات المتوازية مباشرة في المعالجة الآليّة للغة العربية. إذ تستلهم هذه الدراسة تصورها من المفاهيم السياقية اللغوية التي تقدمها المدونات المتوازية، وما توفره من سياقات ترجمية من شأنها أن تمكن الحاسوب من التعرف بشكل أكبر على الدلالات العربية وشحناتها عبر السياق الترجمي. أي أن هذه المدونات المتوازية، الأجنبية–العربية والعربية–الأجنبية، تحمل ما تحمله من دلالات تولدت عبر السياق الترجمي، وتقف هذه الدراسة عند الاستثمار في الدلالات العربية منها، في محاولة لجعل الحاسوب يستوعب أكبر قدر من السياقات العربية المختلفة سواء أكانت عامة أو متخصصة. ويُسهم هذا النهج في تطوير المعالجة الآليّة للغة العربية عبر توفير ما أمكن من هذه المدونات المتوازية وادخالها في البرمجيات الحاسوبية عبر ما يُسمى بذاكرة الترجمة، ومن ثم اغراق هذه البرمجيات في كم هائل من السياقات اللغوية والترجمية تمكنها أولا من التعرف على السياق العربي بشكل أسرع وفي مجالات مختلفة. وثانيا تضاعف فرص توليد السياق العربي آليّا عبر ذلك التوازي الحاصل في السطر البرمجي الواحد للدلالات العربية – الأجنبية المتقابلة.

2.مدخل إلى اللسانيات الحاسوبية:

1.2 ماهية اللسانيات الحاسوبية:

تُعد اللسانيات الحاسوبية من التخصصات البينية فهي تندرج ضمن كل من اللسانيات التطبيقية (Applied Linguistics) وعلم الحاسوب (Computer Sciences) ومعالجة اللغات الطبيعية (Natural Languages Processing) فضلا عن الترجمة الآليّة (Machine Translation)، فهي بذلك تنطوي تحتها ظواهر لغوية وغير لغوية ترتبط ارتباطا وثيقا بمعالجتها إما لغويا أو نُظميا (خوارزميا) (Baldwin & Kordoni, 2011, p. 3).

وتُعنى اللسانيات الحاسوبية بدراسة أنظمة الحاسوب لتمكينه من فهم اللغات الطبيعية وتوليدها، أي أنها تُعنى بمعالجة اللغات الطبيعية في شتى المجالات عبر الحاسوب (علي، 1988، صفحة 5). وقد برز هذا الفرع المتطور بشكل أكبر من المظلة الكبرى للسانيات التطبيقية التي شهدت بدورها تطورا ملحوظا في معالجتها وطرق دراستها مع انتصاف القرن العشرين وبروز الأنظمة الحاسوبية علما واعدا. فازدهرت اللسانيات الحاسوبية في ظل الحاجة الملحة آنذاك إلى حوسبة اللغة قصد السعي إلى محاكاة التطور الحاصل في علم الألة والتكنولوجيا بصفة أخص. ولعل ما سرع في ظهورها وتطورها كفرع نظري متميز عما سبقه من فروع اللسانيات التطبيقية هو ارتباطه المباشر بالترجمة الآليّة، والتي ظهرت بدورها على الساحة بوصفها بديلا (أو معينا) للترجمة البشرية التقليدية في ظل الزخم المتصاعد للنصوص العلمية والتكنولوجية مواكبةً لتطور المجالين (علي، 1988، صفحة 8). في هذا السياق، يبرز تعريف قدمه نيوقس للسانيات الحاسوبية، إذ يقدمها على أنها "فرع علمي يربط اللغة والحاسوب ويهدف إلى تصميم نماذج رياضية للتركيب اللغوي

للتمكن من معالجة اللغة آليّا" (الغامدي، 2017، صفحة 5)، حيث يرى فيها وسيلة فاعلة لتطوير نظرية لغوية جديدة بالاستعانة بالحاسوب. فالأمر إذا لا يقتصر على مجرد حوسبة اللغات الطبيعية (أي برمجتها في الحاسوب)، بل يتعداه إلى تطوير النظام التركيبي لنظم اللغة وتجاوز ما ألفه الباحثون اللغويون من مستويات (صرفية وتركيبية وصوتية وما إلى ذلك). فما نشهده حاليا من نصوص لم يعد يقتصر مداها على الحرف والكلمة والجملة، بل أُدرجت عناصر غير لغوية من صور ومقاطع فيديو ومقاطع صوتية باتت تشكل تأثيرا أكبر لدى القارئ/المتلقي، وهو ما يدعم فكرة تطوير اللغة عبر الحاسوب.

2.2 آلية عمل اللسانيات الحاسوبية:

تقوم اللسانيات الحاسوبية في معالجتها للغة على محاولة تعريف الحاسوب بالعملية الذهنية التي يؤديها العقل البشري في استقباله للغة وانتاجها. وكذا محاولة إدخالها في خوارزمياته الأساسية بحيث يسهل عليه فهم الرموز اللغوية وادراك متطلبات صفّها في الكلمة والجملة والنص كل على حدى (علي، 1988، صفحة 17)، وهو ما يتطلب توصيفا بشريا دقيقا لدى برمجة كل تلك الرموز بما تحمله من ظواهر لغوية قدر الإمكان. أما منتهى الغاية من كل هذه العملية التفاعلية بين العقل البشري وآلية الحاسوب فتكمن في توفير كفاية لغوية للحاسوب (Competence) كما اصطلحها نعوم تشومسكي Naom Chomsky، بحيث تقربه إلى ما يتطلبه العقل البشري من أُسس يبني عليها مداركه للغة ومدى فهمه لها. غير أن هذه الغاية لا تخلو من التعقيد في عمليتها، فتحقيق كفاية لغوية للحاسوب يتطلب في المقام الأول عقلا بشريا مدركا تمام الادراك لحيثيات اللغة محل البرمجة (الحوسبة) ومن ثم تسلسلا توصيفيا (نقليا) لنظم تلك اللغة محل النقل من صرفيات وتركيبات وصوتيات ومدلولات وألفاظ وقواعد وأحكام بيانية وأساليب، وغيرها من المكونات الأساسية لأي لغة طبيعية (العناتي، 2019، صفحة 7). ولعل ما يزيد الأمر خصوصية هو ذلك الاختلاف النوعي الذي نجده بين اللغات على اختلاف انتماءاتها الجغرافية والثقافية والنُظمية على وجه التحديد، فاللغة العربية على سبيل المثال تختلف اختلافا جوهريا في نُظمها عن الإنكليزية مثلا (اللغة الرائدة في المجال). بدءاً باتجاه القراءة والكتابة إلى اختلاف الحرف وبنية الجمل والتراكيب والصرف وأسلوب الطرح وغيرها من الاختلافات. وهو ما يعني عدم تماثل النُهج في عملية حوسبة اللغات بشكل عام. فالنهج المطبق على اللغة الإنكليزية لا يمكن أن يطبق بذاته على اللغة العربية، ففي حال ذلك سينتج لدينا نظاما لغويا غير متوازن في جميع مستوياته، ما يُخل بدوره بسبل فهم اللغة العربية عبر الحاسوب ويجعل من توليدها حاسوبيا أمرا مستعصيا.

إلا أن المشكلة تتجاوز الآليّة إلى مواكبة اللغات الطبيعية كافة لهذه العملية الحوسبية. وعليه، يبرز السؤال حول قدرة اللغات الطبيعية جميعها على مواكبة التطور الحوسبي الحاصل.

الإجابة الأولية ستكون أقرب إلى الإيجابية بالقول "نعم"، فجل اللغات الطبيعية تستند أُسسها على المستويات ذاتها وهي بالتالي تبقى قابلة للمعالجة الآليّة عبر الحاسوب، فكيف نبرر إذا التفاوت الحاصل بين اللغات بهذا الصدد؟

لعل المتابع للتطور الحوسبي لأي لغة كانت سيجده حتما مرتبطا بمدى تطور الناطقين بها (ثقافيا، معرفيا وخاصة علميا وتقنيا). فبعد التطور العلمي والتكنولوجي الذي شهدته أوروبا على أعقاب الحرب العالمية الثانية، وتحديدا بين العامين 1955 و 1965، حيث قاد باحثون أمريكيون وبريطانيون مشاريع تهدف إلى حوسبة العمليات الترجمية واضفاء طابع أكثر ميكانيكية عليها لاسيما من اللغة الروسية (Britannica, 2021, p. 1). وعلى اثر ذلك، عرفت اللغات الأوروبية تطورا ملحوظا على مستوى المعالجة الآليّة، فهيمنت اللغة تلك اللغات على عالم اللسانيات الحاسوبية منذ استهلالها، بينما لم تحظ لغتنا العربية بالسعي والاهتمام ذاتهما فتعسر تطورها في هذا المجال. وظل الأمر على حاله على الرغم من بعض المحاولات من هنا وهناك على المستوى النظري من قبيل الدكتور نبيل علي ونموذجه في ادخال اللغة العربية إلى الحاسوب. إلا أنه وبشكل عملي أُدرجت اللغة العربية لأول مرة في القوائم الحاسوبية سنة 1986، عندما قام الباحث الجزائري المعروف بشير حليمي بتعريب الحاسب الألي ليتمكن العرب من استخدام الحاسوب بالخط العربي بالطريقة التي نشهدها حاليا. إلا أنه ورغم ذلك، لم تحظ اللغة العربية بما قد نسميه ثورة في المجال وظل الامر محدودا حتى شيوع الحاسوب وأفرعه اللسانية علما وانتشارا مع نهاية التسعينات وبداية الألفية الجديدة، حيث لُوحظ التحسن الكبير في المعالجة الآليّة للغة العربية عبر الحاسوب.

وقد أُعتمد في سبيل تطوير حوسبة اللغات نُهج عديدة بانتماءات مختلفة. إلا أن اعتماد لغات جمة على المدونات باختلافاتها النمطية، لاسيما ما توازى منها بلغتين أو أكثر، بات يمثل نهجا جديدا آخذ في الانتشار في الآونة الأخيرة، وصار يحظى باهتمام الباحثين في الميادين ذات الصلة باللغات وبالعلوم الحاسوبية وما بينهما. فكيف يسهم هذا النمط المتوازي من المدونات في حوسبة اللغات؟ وما مدى استثماره في تطوير الحوسبة العربية؟

3.المدونات المتوازية أداة ناجعة في تطوير الحوسبة العربية:

1.3 ماهية المدونات المتوازية:

يُعرف ليتش Leech –بترجمة تلخيصية– المدونة بأنها قواعد بيانات (Data Bases) تنظيمية ضخمة لعدد من النصوص الالكترونية التي صُممت وفق معايير معينة لتُستخدم في البحث اللغوي (ليتش، 2010، صفحة 4). وتتفرع أنماط المدونات بحسب التعدد اللغوي فيها، فمنها أحادي اللغة (Monolingual Corpora) نجدها تعالج مواضيع جمة بلغة واحدة، بحيث يمثل هذا النمط النهج التقليدي للمدونات على اختلافها. وقد تتعدد اللغات في المدونة الواحدة (Multilingual Corpora) فنجدها في غالب الأحيان تعتمد على نص واحد أصلي وباقي النصوص ترجمات مباشرة له بلغات متعددة. ويحدث أن ترد المدونة بلغتين مختلفتين (Bilingual Corpora) يتوازى فيها نصان يناقشان موضوعا واحدا كل بلغته (Parallel Corpora)، فإما ان يكون احداهما ترجمة للآخر أو أن يكونا نصين أصليين في اللغة التي كُتب كل منهما بها. وتركز هذه الدراسة على نمط توازي النصوص في حد ذاته من خلال ما يسمى بالمدونات المتوازية.

ويعرف سينكلير Sinclair –بترجمة تلخيصية– المدونات المتوازية بأنها "نصوص أصلية تقابلها ترجمات بلغة أو بلغات أخرى" (1991، صفحة 19)، في حين تعرفها منى بايكر Baker بأنها "مجموعة نصوص أصلية كُتبت بلغة الأصل (أ) تقابلها ترجماتها باللغة (ب)" (1995، صفحة 235).

وقد لاقى هذا النوع من المدونات رواجا كبيرا في مجال اللسانيات التطبيقية ودراسات الترجمة، فضلا عن اللسانيات المقارنة وعلم الاصطلاح، وبات مؤخرا يشغل العديد من الدراسات في علم اللسانيات الحاسوبية والترجمة الآليّة على وجه الخصوص. وقد اكتسبت المدونات المتوازية أهمية كبيرة فيما تقدمه من تصورات جديدة حول المعالجة السياقية للغتين (الأصل والترجمة) في المدونة المتوازية الواحدة، وهو ما لا توفره المدونات التقليدية أحادية اللغة. حيث مكنت الباحثين من الاستثمار في تلك المقارنة بين النصين على مستويات عدة، فمكنتهم من التعرف على الفروق الكامنة بين اللغات محل التوازي في تلك الخصال الدلالية الثقافية والسياقية. فضلا عن العناصر التركيبية واللغوية التي تعتمدها اللغات في تناولها للمواضيع ذاتها ومدى توافق سياقاتها في معالجتها.

2.3 استخدامات المدونات المتوازية في الترجمة الآلية:

اتجه الباحثون والمهتمون بالترجمة الآليّة حديثا إلى اعتماد المدونات المتوازية في المعالجة الآليّة للبرمجيات الحاسوبية الحديثة. إذ يسعى هؤلاء إلى فسح المجال نحو جمع أكبر عدد من هذه المدونات على الشبكة العنكبوتية لإيمانهم الراسخ بمدى الاسهام الذي تضفيه في تطوير المعالجة الآليّة لهذه البرمجيات أثناء الترجمة. ما يمكّن من الحصول على نتائج أفضل للترجمات الآليّة، التي طالما يُعاب عليها قصورا في جودتها، لاسيما إذا ما تعلق الأمر بلغتين بالغتي الاختلاف في النُظم كما هو حال اللغة العربية والإنكليزية مثلا.

وعلى الرغم من كل هذا الاهتمام المتصاعد بالمدونات المتوازية في عموم الأبحاث العالمية، إلا ان اللغة العربية لم تحظ بالاهتمام المطلوب. فهي لا تزال تفتقر إلى الجهود الرامية إلى توفير مصادر لغوية مماثلة. وهو ما عسر من مهمة اللغة العربية في مواكبة التطور الحاصل على هذا المستوى والذي تقوده اللغات العالمية الكبرى، وباتت تحقق فيه سبقا واضحا. ويُعزي معظم الباحثين العرب هذا التخلف إلى الصعوبة التي تكتنف العملية الحوسبية للغة العربية، والتي لا تزال تشكل قيودا على الرغم من التطور الملحوظ مؤخرا في مدى تعرف الحاسوب على نُظم اللغة العربية اللغوية، إذ بتنا قادرين على ملاحظة هذا التقدم مرأى العين مختصين وعوام. لذا، أصبح التوجه اليوم مرتكزا على تعريف برمجيات الحاسوب بالدلالات العربية وسياقاتها المختلفة في المجالات العامة والمتخصصة، قصد تمكينها من تحليلها ومعالجتها باستمرار بغية الوصول إلى المرحلة التوليدية لها. أي أن يتمكن الحاسوب من استيعاب الدلالات العربية على اختلاف انتماءاتها السياقية، فيتمكن على إثر ذلك من توليدها في العمليات الترجمية اللاحقة.

وتنطلق هذه الدراسة من هذا المسعى تحديدا، من خلال محاولة الاستثمار في السياقات اللغوية التي توفرها المدونات المتوازية العربية–الأجنبية والأجنبية–العربية. والاستثمار فيما يتمخض عنها من سياقات تُرجمية من شانها الاسهام

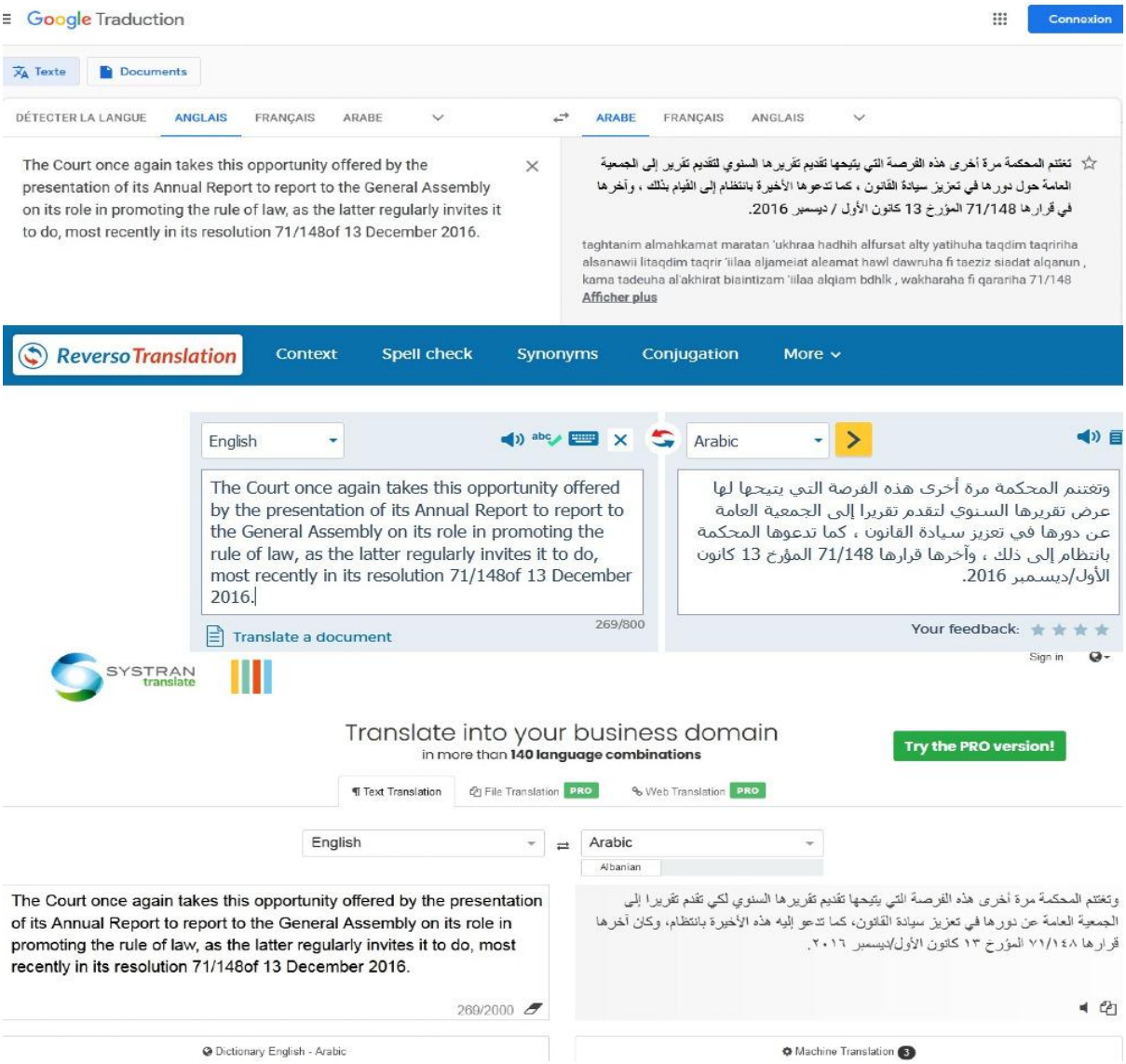
في تزويد الحاسوب بالدلالات العربية المختلفة في الميادين عامها ومتخصصها، وما تلازمها من مصطلحات حاملة لهذه الدلالات والشحنات المختلفة.

4. النهج السياقي الموازي واسهامه في تطوير الحوسبة العربية:

عند توازي اللفظ (أو المصطلح) وما يقابله لغويا واصطلاحيا بلغة أخرى في النص الموازي الواحد، فإنه يتولد لدينا ما يسمى بالسياق الترجمي. وهو ما سيستخدمه الحاسوب فيما بعد (بالاستعانة بنُظم الذكاء الاصطناعي) في العملية الترجمية للنصوص التي تحمل السياق الترجمي ذاته (Zanettin, 2011, p. 5). وتستمر العملية بهذا الشكل بحيث يخزن البرنامج الآلي هذه السياقات في برمجياته لتساعده في التعرف على السياقات الجديدة المدرجة والتي تتماثل مع ما خزنه من مجموعة النصوص الموازية المدرجة مسبقا، فتتعدى العملية مرحلة المقابلة والمطابقة لتصل إلى مرحلة التوليد التفاعلي (تفاعل الألة مع ما خزنته وما تستقبله من نصوص) للسياقات المختلفة، كل حسب مجال انتمائه، فيمكنها ذلك الذكاء الاصطناعي من المعالجة السريعة من جهة، ومن التعرف على الأسلوب المتماهي في هذا السياق الجديد (أدبي، علمي، شاعري ...إلخ) من جهة أخرى.

في هذا السياق، تعكف هيئة الأُمم المتحدة على اتاحة تقاريرها بلغاتها الست لتُستخدم في برامج الترجمة الآليّة وخدماتها عبر الانترنت، فتستخدمها هذه البرامج على أنها نصوص موازية عند الترجمة بين اللغات المدرجة (العربية، الانكليزية، الفرنسية، الاسبانية، الصينية والروسية). حيث تترجم هذه التقارير إلى اللغات الخمس من اللغة الانكليزية بشريا، وتحملها في موقعها الإلكتروني لتتيحها بذلك للاطلاع الآلي. وعند إدراج أي من هذه النصوص في من تلك البرامج أو الخدمات الترجمية عبر الانترنت، تقوم هذه الخدمة أو هذا البرنامج للترجمة الآليّة باستعراض النص الأخر (المقابل والمترجم بشريا) في اللغة المستهدفة (من بين اللغات الست). في هذه الدراسة، اعتمدنا أداء ثلاث خدمات ترجمية عبر الانترنت (Google Translate وReverso وSystran Translate) كونها ضمن الخدمات الترجمية التي تعتمد برمجياتها على نُظم الترجمة الالية العصبية، والمستندة أساسا على المدونات المتوازية التي تُرجمت نصوصها بشريا. فضلا عن اتاحتها للاستخدام المجاني ما يعزز انتشارها بين عموم المستخدمين. و قد أدرجنا هذا المقتطف الآتي (من تقرير للأمم المتحدة، رابطه: https://undocs.org/en/A/72/4) لإيضاح المقصد:

الشكل (1): تماثل الأداء الترجمي للمقتطف الإنكليزي ذاته في ثلاث خدمات للترجمة الآليّة عبر الانترنت



المصدر: تقرير محكمة العدل الدولية، 2016، ص 9

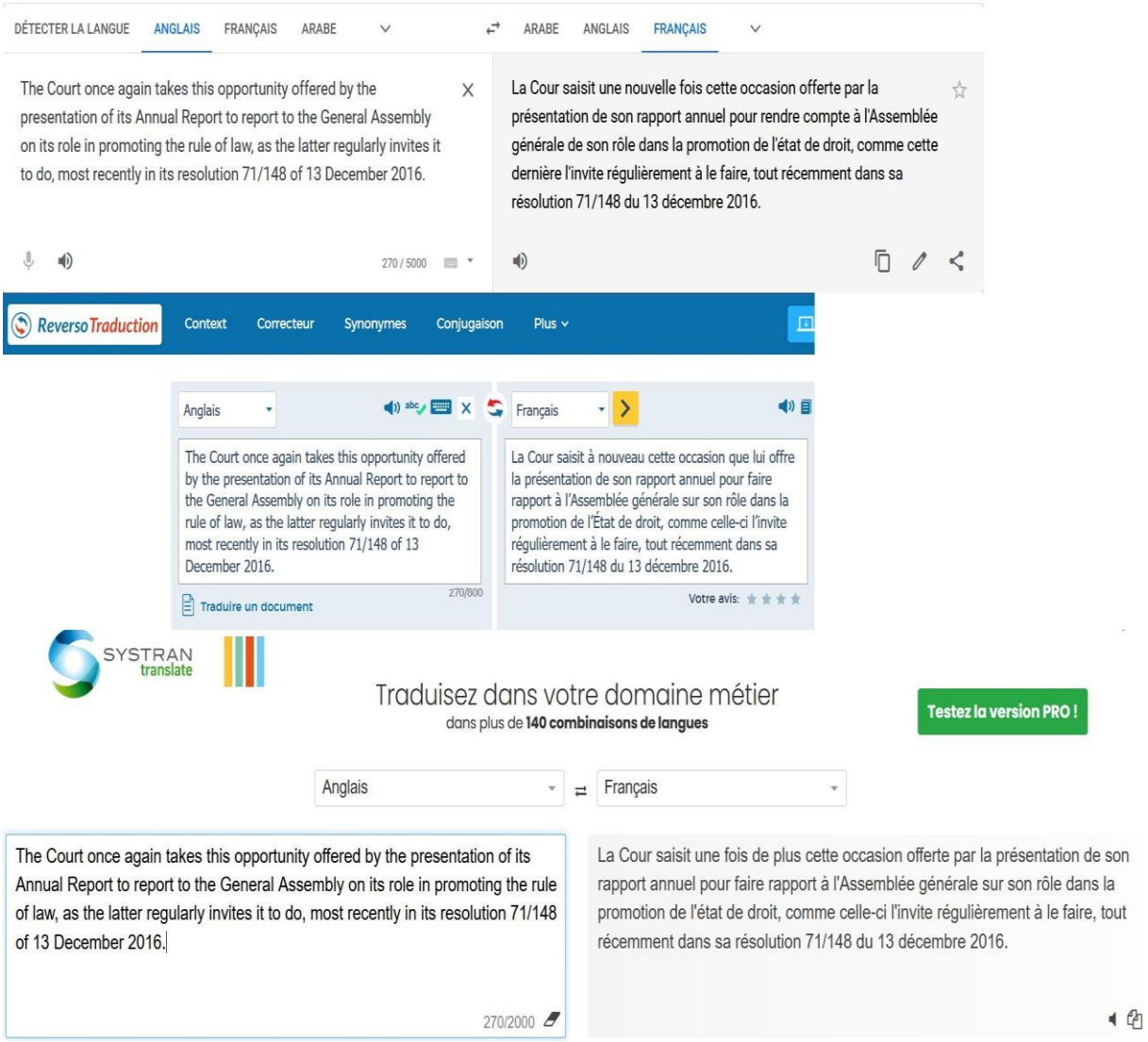
فكما هو ظاهر من الشكل 1، أدرجنا مقتطفا من تقرير لهيئة الأمم المتحدة (مرجعه في الأعلى) في ثلاث خدمات للترجمة الآليّة عبر الأنترنت (Google Translate وReverso وSystran Translate)، حيث تبين تماثل الأداء الترجمي في الخدمات الثلاث، وهذا راجع لاعتمادها على تلك التقارير بصفتها ذاكرة للترجمة حيث تستخدمها في هذه الحال كمدونة متوازية خلال العملية وإن لم تتوازى ظاهريا، لكنها تتوازى في برمجياتها لتستخرج عل اثر ذلك النص المكافئ في ثوان معدودات.

وما يؤكد ذلك، أن هذه الخدمات الثلاث وبالرغم من اختلاف خوارزمياتها وبرمجياتها، إلا أنها استطاعت الوصول إلى السياق الترجمي المقصود بسهولة كبيرة، فتساوت نتائجها جميعها وإن اختلفت صياغتها في أخر المقتطف. وهي حقيقة تثبت مدى اسهام المدونات المتوازية في تسهيل الوصول إلى السياقات الترجمية، وبالتالي الوصول إلى السياقين اللغويين

للغتين محل العملية الترجمية. وهو ما يمكن هذه البرمجيات من التعرف بشكل أكبر على اللغات المعنية في الطرح السياقي وثم فسح المجال لتطويرها آليّا وحوسبيا من خلال سياقاتها الواردة في توازي النصوص المدرجة بها.

على صعيد أخر، أدرجنا المقتطف ذاته في الخدمات الترجمية الثلاث، لكن هذه المرة بالاعتماد على اللغتين الانجليزية (المصدر) والفرنسية (الهدف)، إذ نتج ما يلي:

الشكل (2): تطابق الأداء الترجمي للمقتطف الإنكليزي إلى اللغة الفرنسية في الخدمات الترجمية الثلاث

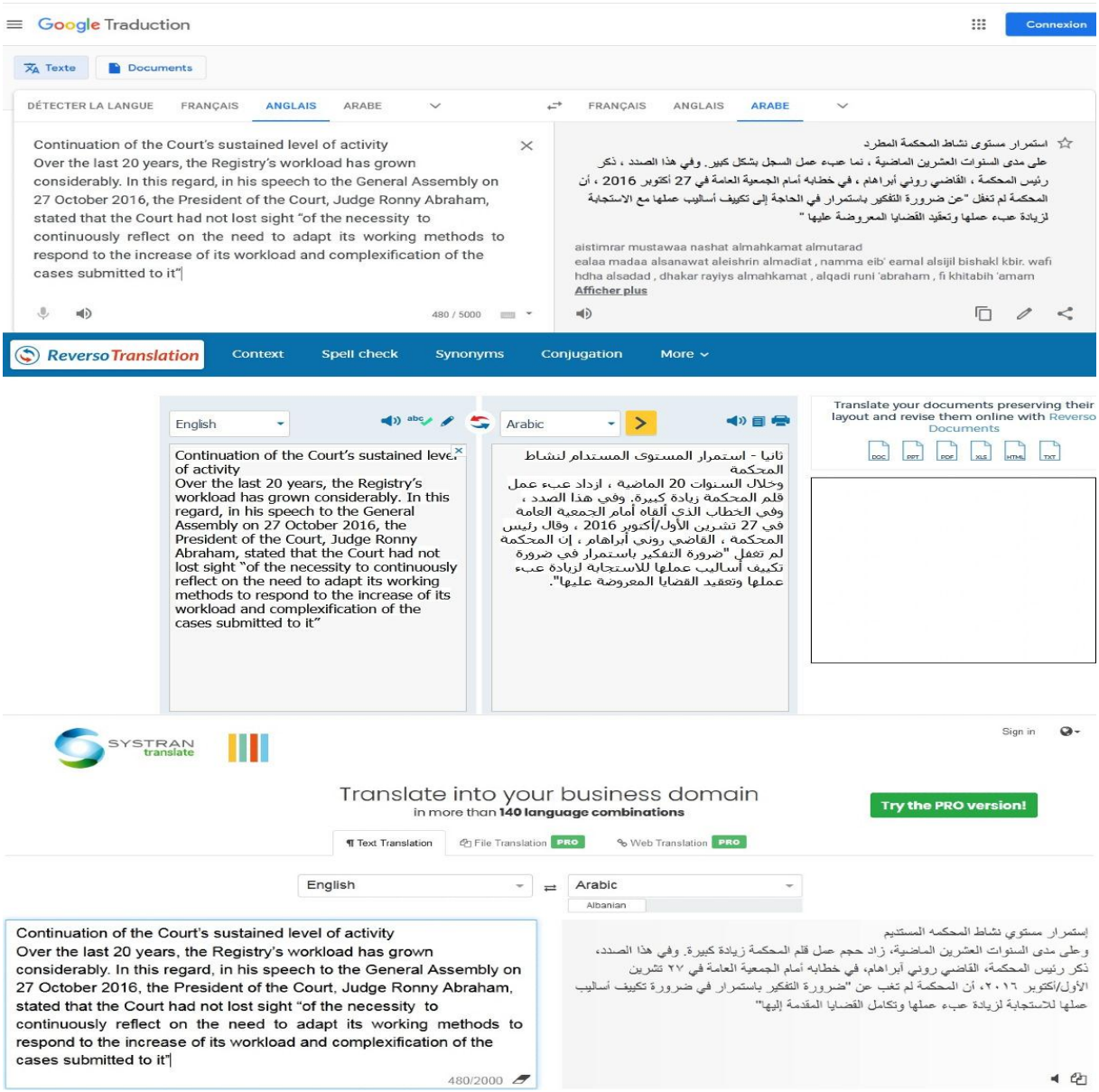


المصدر: تقرير محكمة العدل الدولية، 2016، صفحة 9

ما يتضح من أداء الخدمات الثلاث لهذا المقتطف من اللغة الانكليزية إلى الفرنسية، أن هذه الخدمات قد اختلف أداؤها الأسلوبي لعدد من الجمل. لكن ذلك لم يؤثر على المعنى العام لها، بل عزز أداء برمجياتها كونها لم تطابق مفرداتها بعضها البعض، بل عالجت السياق وولدت على اثره ما يكافئه في اللغة الهدف.

بينما أدرجنا مقتطفا أخر من سياق مشابه في الخدمات الترجمية ذاتها، إذ تباينت النتائج على عكس المثال الأول. حيث أدرجنا مقتطفا من تقرير الأمم المتحدة بشأن قرارات اتخذتها محكمة العدل الدولية (رابطه: https://undocs.org/en/A/72/4):

الشكل (3): توضيحي لعدم تماثل الترجمة إلى العربية من المقتطف الإنكليزي ذاته في الخدمات الترجمية الثلاث



المصدر: تقرير محكمة العدل الدولية، 2016، صفحة 8

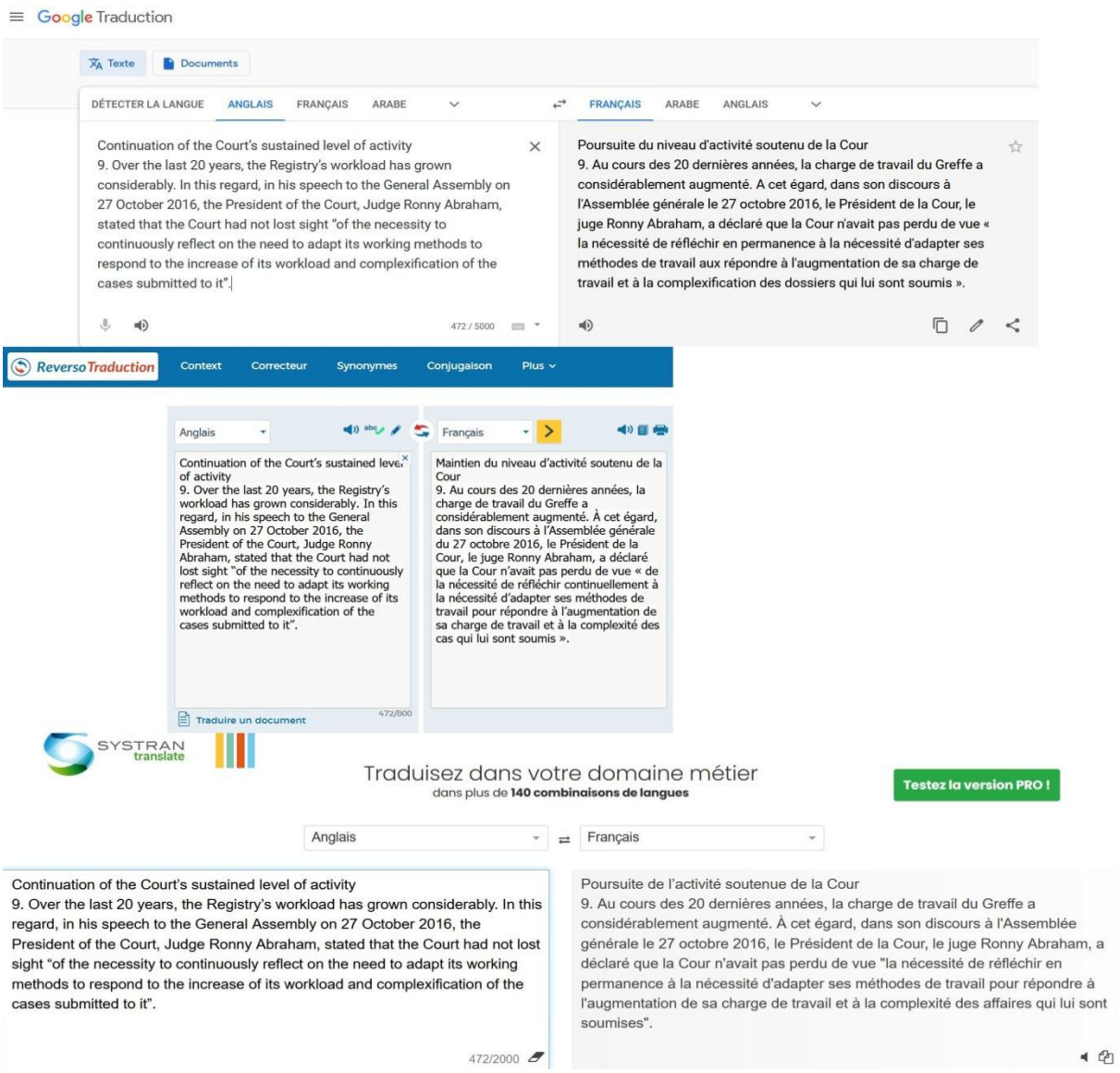
ما يُلاحظ من ترجمة هذا المقتطف هو عدم التماثل فيها على عكس المثال الأول، على الرغم من كون هذا التقرير مُدرج كذلك باللغات الست للأمم المتحدة حاله حال مقتطف التقرير الأول. لكن ما يلاحظ كذلك هو ثبات السياق، فذلك النهج السياقي مكن برمجيات هذه الخدمات من الوصول إلى المصطلحات المكافئة واللازمة في السياق، ومن ثم

عبرت هذه الخدمات الترجمية الثلاث عن السياق كل بأسلوب مختلف. وهو ما يقر بفعالية النهج السياقي الذي تضفيه المدونات المتوازية في تعزيز جودة الترجمة، وفي تقريب الدلالات السياقية العربية إلى البرمجيات الحاسوبية.

لكن ما الذي يبرر عدم تماثل النتائج في هذا المقتطف والذي تُرجم كنظيره في المثال الأول إلى اللغات الخمس؟

تفسير ذلك حسب الدراسة، يعود إلى السياق، أي أن هذا السياق في هذا النوع من التقارير شحيح في برمجيات هذه الخدمات خاصة باللغة العربية. وبالتالي تداخلت الدلالات وعجزت هذه البرمجيات عن تحديد المقصود منها، ونتج ما نتج من تفاوتات في الأداء الترجمي (على عكس ما أظهره أداؤها لذات المقتطف إلى اللغة الفرنسية، أنظر الشكل 4 أدناه)، وعلى عكس المثال الأول الذي تشربت برمجيات هذه الخدمات من سياقاته العربية وبات هينا عليها استخراج الدلالات المكافئة، وهو ما أفرز نصا مطابقا في الخدمات التُرجمية الثلاث، أي أن شُح المدونات المتوازية التي تصب في هذا السياق عسر على هذه البرمجيات التعرف على الدلالات اللازمة في السياق العربي دونما تداخل أو لُبس.

الشكل (4): يوضح تماثل ترجمة المقتطف الإنكليزي ذاته إلى اللغة الفرنسية بين الخدمات الترجمية الثلاث



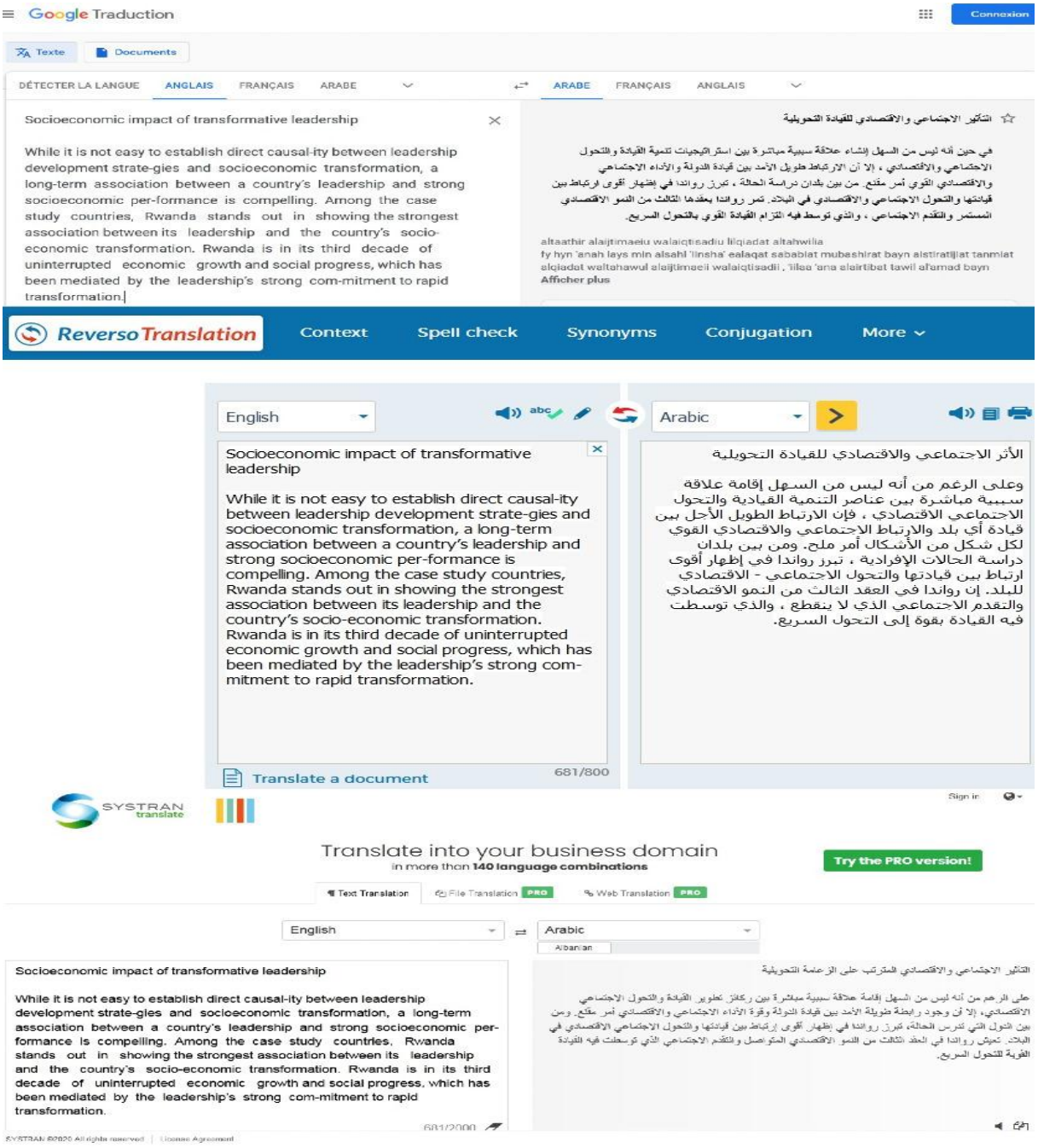
المصدر: تقرير محكمة العدل الدولية، 2016، صفحة 8

وحل ذلك حسب ما تقترح هذه الدراسة يكمن في تزويد الحوسبة الآليّة وبرمجياتها بكل أنواع النصوص المختلفة وذات السياقات المتعددة في شكل مدونات متوازية، لتعريف الحاسوب أكثر بكيفية معالجة اللغة العربية لدلالاتها مقارنة بنص اللغة الموازية، ويتمكن على اثر ذلك من استخراجها وتوليدها في السياقات المماثلة.

ويسهم هذا النهج في تطوير الذاكرة الترجمية لدى البرمجيات الآليّة والحاسوبية بشكل كبير، خاصة إذا ما تعلق الأمر باللغة العربية التي لا تزال تشق طريقها نحو الوصول إلى الجودة المرجوة مقارنة بباقي اللغات الخمس التي أحرزت تقدما واضحا وباتت تسبق اللغة العربية في هذا المجال سبقا جليا. غير أن هذا لا يقتصر على تطوير الترجمة فحسب، بل يسهم في تطوير الحوسبة العربية من خلال السياق، فالبرمجيات الآليّة وفق هذا النهج ستتمكن من التعرف أكثر على السياق العربي وترفع من مستوى معالجتها إليه، فتتجاوز على إثر ذلك الصعوبات الكامنة في المستويات اللغوية الصرفة والتي شكلت وتشكل التباسات متعددة أسهمت بدورها في تأخر تطوير اللغة العربية عبر الحاسوب. ما نطرحه هنا، أن الوحدة الترجمية عند الترجمة من اللغة العربية وإليها خاصة ستتجاوز ما ألفته البرمجيات الحاسوبية والآليّة بعمومها، لتعتمد السياق مباشرة في العملية الترجمية، ولن يتأتى ذلك إلا من خلال تزويدها بالمدونات المتوازية بالأخص قدر الامكان، فبالقدر الذي تستوعبه هذه البرمجيات من سياقات عربية مختلفة ومتنوعة بالقدر الذي تتمكن فيه بشكل فعلي من الوصول إلى الفهم الوافي والمعالجة الدقيقة للسياق العربي فيما تقوم به من عمليات توليدية للغة العربية.

وتظهر أهمية هذا الاسهام جلية في الميادين المتخصصة على اختلافها، ففي الحين الذي تهيمن فيه العولمة على ربوع الأرض جميعها، برزت سياقات متخصصة جمة باتت بدورها تطرح مشكلات ترجمية ولغوية عديدة. حيث نشهد حاليا تطورا هائلا في علم المصطلح المتخصص وباتت تُقاس درجة تطور لغة ما وفق ما تحمله من شحنات متخصصة تصبها في مصطلحات من توليدها، لتستخدم في سياقات بعينها دون غيرها. وهو ما يعزز فكرة اعتماد السياق اللغوي كوحدة ترجمية عبر الاستثمار في السياقات التي تقدمها المدونات المتوازية قصد السعي إلى تحقيق التخصص في اللغة العربية، ما يعني أن نستثمر في الجهود البشرية عبر ما يقدمه المترجمون من ترجمات في السياقات المتخصصة على وجه التحديد، لندرجها في البرمجيات الآليّة تماما كما أوضحنا في المثال التوضيحي أعلاه. ولتبيان مدى أهمية الاستثمار في اعتماد المدونات المتوازية في محاولة للنهوض بالعملية الآليّة لحوسبة اللغة العربية، نستعرض في المثال الاتي مقتطفا من تقرير للاتحاد الافريقي (الذي لا يعتمد على ترجمة تقاريره إلى اللغات الرسمية لديه على عكس هيئة الأمم المتحدة) حول تعزيز القيادة التحويلية لتنمية إفريقيا (2019)، حيث أدرجنا المقتطف في خدمات الترجمة الآليّة الثلاث ذاتها التي استخدمناها في المثالين الأول والثاني، إذ نتج ما يلي:

الشكل (5): توضيحي لكيفية أداء الخدمات الترجمية الثلاث إلى اللغة العربية في ظل غياب مدونة متوازية



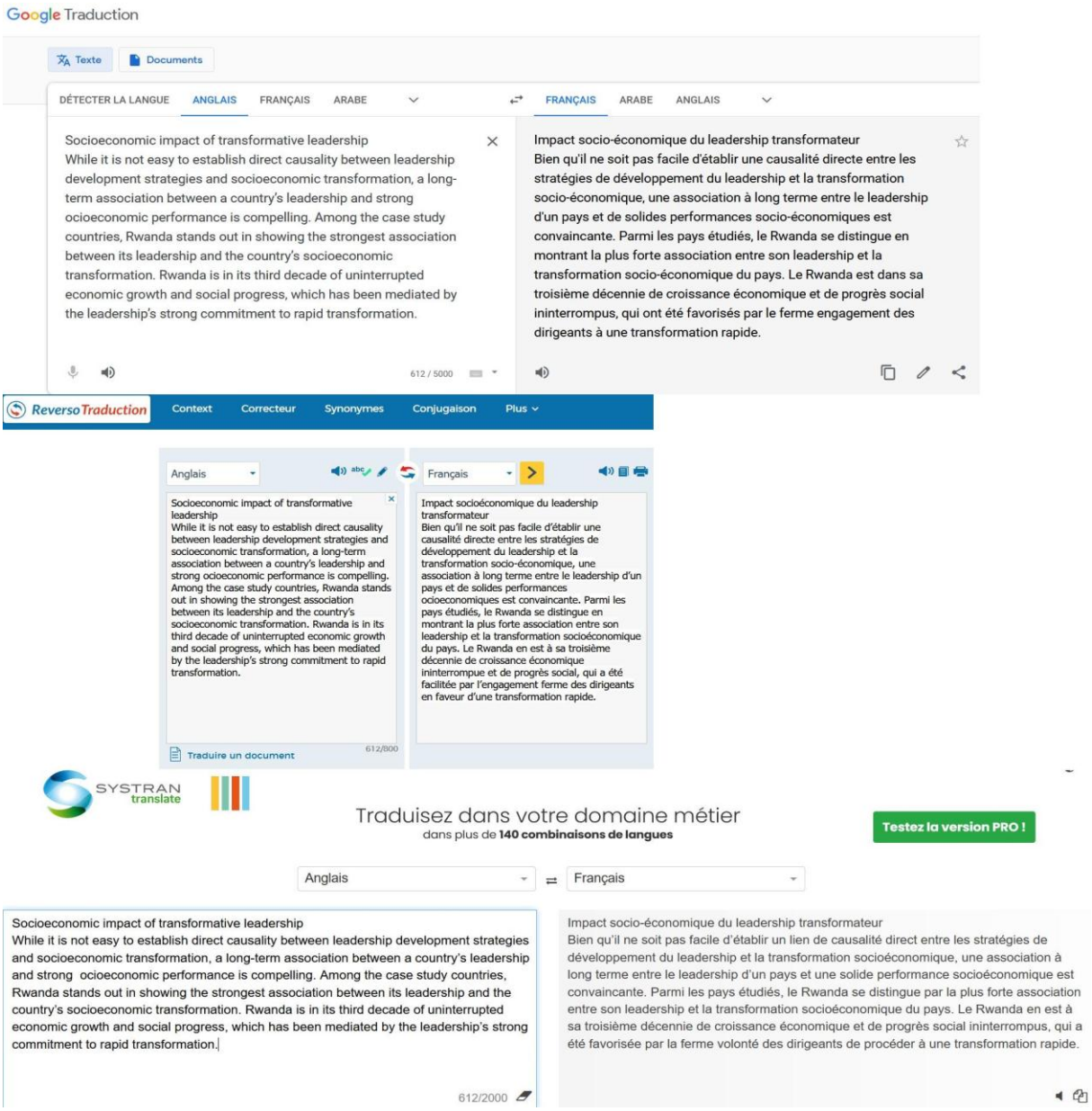
المصدر: تقرير القدرة الافريقية (عن الاتحاد الافريقي)، 2019، صفحة 9

مما يتضح من الشكل التوضيحي أعلاه، ذلك التباين المتدرج بين كبير ونوعي في السياق في حد ذاته إضافة إلى عدم ثبات المصطلحات المقابلة والتي تؤدي دورا محوريا في نسغ السياق العام لهذا النوع من التقارير، فمثلا نجد مصطلح Leadership يُترجم تارة إلى "قيادة" وإلى "زعامة" تارة أخرى، وهما مصطلحان يختلفان جذريا في أساليب استخدامهما في اللغة العربية، ومصطلح Strategies يُترجم بدوره إلى ثلاث صيغ مختلفة: "استراتيجيات"، "عناصر" و"روابط"، وهي صيغ غير قابلة للمبادلة في هذا السياق بالتحديد. فضلا عن بعض التراكيب الغريبة على اللغة العربية

والمتلازمات اللفظية البعيدة عن بنية اللغة العربية واستخداماتها اللازمة والمتأصلة في النظام اللغوي العربي، لاسيما منها ما يرتبط بالسياق مباشرة. فعندما افتقرت برمجيات هذه الخدمات إلى مدونات متوازية تصب في السياق ذاته نجم هذا التباين الذي من شأنه الاضرار أولا من جودة الحوسبة العربية ومسار تطورها في هذا المجال. وثانيا يسهم بشكل أكبر في إضفاء التباسات إضافية للمصطلح العربي في مختلف السياقات المتخصصة على وجه التحديد.

وتبيانا لواقع حال حوسبة اللغة العربية دلاليا مقارنة باللغات المنافسة، أدرجنا المقتطف ذاته في الخدمات الترجمية ذاتها قصد ترجمته إلى اللغة الفرنسية كما يوضح الشكل الآتي:

الشكل (6): توضيحي لكيفية أداء الخدمات الترجمية الثلاث للمقتطف إلى اللغة الفرنسية في ظل غياب مدونة متوازية



المصدر: تقرير القدرة الافريقية (عن الاتحاد الافريقي)، 2019، ص 9

ما يتضح من أداء برمجيات هذه الخدمات الثلاث، هو الثبات المصطلحي الكبير بل والمتطابق في كل منها. وهو ما نتج عنه ثبات سياقي تام للنص المترجم مقارنة بالنص الأصل. فالمصطلحات التي تأرجح ناتج ترجمتها فإلى اللغة العربية، حافظت على شحنتها كاملة في الترجمة الفرنسية. فتُرجم مصطلح "Leadership" إلى "Leadership" في الخدمات الثلاث، وكذلك الحال مع مصطلح "Strategies" والذي تُرجم إلى "Stratégies". ويعود ذلك إلى مدى تعرف برمجيات هذه الخدمات الترجمية على الدلالات الفرنسية بشكل أكبر مقارنة بنظيراتها في اللغة العربية. ما نحاول الاشارة إليه في هذه النقطة، أن توليد برمجيات هذه الخدمات للدلالات السياقية الفرنسية على سبيل المثال أيسر لديها من توليدها للدلالات السياقية العربية. وهو ما يشير إلى شح السياق العربي فيها بشكل عام، وافتقار برمجيات هذه الخدمات إلى الدلالات العربية اللازمة لتوليدها في سياقات مختلفة، دون مساعدة المدونات المتوازية وما تقدمه من سياقات لغوية وترجمية جاهزة.

ونشير هنا إلى أن هذه البرامج تعتمد في فهمها للنصوص المدرجة للترجمة أولا على المستوى اللغوي، حيث تستند عليه في الوصول إلى المعاني التي بني عليها النص المدرج، بحيث تستوعب النص استيعابا كاملا عبر سياقه، وبذلك تصبح قادرة على استيعاب معان جديدة عبر ما أدرج لديها من نصوص تحمل معان مشابهة أو معان مطابقة أحيانا. ثم تمر في المرحلة الثانية إلى السياق الترجمي لتستند إليه في هذه المحاولة التوليدية، من خلال ما أدرج من نص مواز أو بالأحرى من معان موازية في اللغة الأخرى. وبالتالي ينتج لدينا سياق ترجمي جديد يخزن داخل برمجيات الخدمة الآليّة ليستخدم في معالجة النصوص ذات السياق المشابه (Carbonel, 2006, p. 20). لكن الإشكال الوحيد في هذه العملية أن هذه البرمجيات تبقى دائمة التطلب للمدونات المتوازية قدر الإمكان. فبالقدر الذي ندرج فيه تلك المدونات المتوازية إلى النظم التي ينبني عليها هذا البرنامج الآلي، بالقدر الذي يولد نتائج أفضل، ما يعني أن كلما زاد عدد المدونات المتوازية المدرجة في هذه البرامج كلما زاد انتاجها نوعا وجودة.

وتسهم المدونات المتوازية في اطار حوسبة اللغة العربية عبر هذا النهج السياقي في تجاوز المراحل الأولية التي تمر بها المعالجة الحاسوبية للغة (أو النص)، على أساس المستويات التي تنبني عليها اللغة من مستوى صوتي وصرفي وتركيبي وغير ذلك. فيتجاوز الأمر إذا المرحلة البنيوية الصرفة من خلال خوارزميات النقل اللغوي، ليتدرج ليصل في الأخير إلى مرحلة التقابل أو التكافؤ السياقي للنصوص. فالبرنامج الآلي سيجد أمامه نصوصا متوازية تتناول سياقا واحدا بلغتين أو أكثر. وبالتالي يعتمد على هذا السياق في اعادة توليد نصوص مشابهة، بالاستناد على ما اكتسبه من مدخلات أساسية للغة دون الخوض في التفاصيل الدقيقة التي تكتنف اللغة العربية من تنوين ومد وهمز، وما إلى ذلك من الخصوصيات التي تتسم اللغة العربية بها وتختص بها دونا عن جل لغات العالم. وهو ما من شأنه أن يضفي جودة على الناتج النهائي للترجمة، فالوحدة الترجمية هنا ليست لفظا أو كلمة أو جملة أو نصا بل هو السياق، فالسياق هنا هو الذي يولد النص عبر نُظم وبرمجيات هذه الخدمات التُرجمية.

ما سبق يسهم دون شك في تطوير المعالجة الآليّة لترجمة النصوص، لكنه يسهم أيضا بشكل كبير في تطوير حوسبة اللغة العربية. كون ما تنتجه البرامج الآليّة من نصوص تُرجمت إلى اللغة العربية بتلك الجودة التي تفرزها، سوف تخزنها في برمجياتها لاستخدام مماثل فيما بعد. وهو ما يضمن جودة أفضل في سياقات مماثلة على الأقل، خاصة إذا ما تعلق الأمر بالميادين المتخصصة ذات التقانة العالية.

إلا أن كل هذه العملية قد لا تمكن الحوسبة العربية من الوصول إلى المواكبة التامة للتطور الحاصل في المجال عالميا. لأن كل ما سلف ذكره من خطوات في سبيل دمج النهج السياقي في الحوسبة العربية، قد تجاوزته اللغات الأخرى ذات السبق في المجال. وعليه، فإن تطوير حوسبة اللغة العربية يمر حتما عبر الارتقاء بالأليات المتَبعة. في هذا السياق، نشير إلى ما قدمه Ashengo وAga وAbebe في مقال موسوم بـ " Context-based Machine Translation with Recurrent Neural Network for English-Amharic Translation"، حيث قدم الباحثون نموذجا حول سُبل تطوير الترجمة الألية للنصوص من اللغة الانكليزية إلى اللغة الأمهرية (لغة رسمية في الدستور الاثيوبي)، عبر الاستثمار في عدد من الأليات من ضمنها المدونات المتوازية.

في طرحهم للنموذج السياقي الذي تقدمه المدونات المتوازية في سبيل تطوير الترجمة الالية بين اللغتين الانكليزية والأمهرية، يرى الباحثون بأنه يتوجب الاعتماد على مدونات متوازية تتجاوز وحدة الترجمة فيها مستوى المفرد أو المصطلح في حد ذاته. ما يعني أن الترجمة داخل المدونة المتوازية الواحدة يجب أن تقوم على السياق أولا، قبل ادراجها في برمجيات الخدمة أو البرنامج الترجمي (Ashengo, Aga & Abebe, 2021, p.2).

ومن خلال إسقاط هذا التوجه على واقع حوسبة اللغة العربية، نجد بأن تقارير هيئة الأمم المتحدة تكاد تكون المصدر الوحيد الذي يتيح ترجمات بشرية إلى اللغة العربية داخل المدونة المتوازية الواحدة. إلى جانب عدد من المحاولات في بعض أدلة الاستعمال ذات الطابع التقني والعلمي. وهو ما يعد غير كافيا لتحقيق الكفاية الدلالية والسياقية اللازمة لبرمجيات الخدمات الترجمية على الشبكة العنكبوتية، من أجل تعرف أفضل على الدلالات العربية على اختلافها وتوليدها عند الاقتضاء. لاسيما في ظل غياب المدونة المتوازية أداةً مباشرة خلال العملية الترجمية. فاللغة العربية إذا لا تزال تعاني قصورا في الثبات المصطلحي والسياقي حين توليدها من هذه البرمجيات الحاسوبية، لاسيما عند الغياب التام للمدونة المتوازية، تماما كما حدث في المثال الوارد في الشكل (5) أعلاه، أين تأرجح المصطلح عند الترجمة إلى العربية، بينما ظل ثابتا في المثال الوارد في الشكل (6) عند الترجمة إلى اللغة الفرنسية.

هذا القصور الذي تبديه الترجمات العربية على المستوى السياقي، ليس سببه شُح المدونات المتوازية الداعمة للعربية على الانترنت فحسب. أغلب المدونات المتوازية الانكليزية – العربية تعتمد على منهج "صف الجمل" لا "السياقات"، أي أنها ترجمات مباشرة للجمل الانكليزية (أو الأجنبية عموما) إلى اللغة العربية. فالسياق هنا يتولد من ترابط تلك الجمل وفق منهج PBMT (Phrase Based Machine Translation)، الذي يندرج تحت المنهج السياقي للترجمة الألية CBMT (Context Based Machine Translation) الذي جاء به كاربونيل Carbonel (2006) في مقال

بعنوان "Context Based Machine Translation"، حيث يتولد السياق العام للنص من تفاعل الجمل المترابطة منذ البداية إلى النهاية (Popovic, 2017, p. 1). ووظيفة البرمجيات الحاسوبية هنا هي التعرف على ذلك السياق الكامن في ذلك الترابط الجُملي ونقله بذاته إلى اللغات المختلفة، وهو ما يمكنها من التعرف على عدد كبير المفردات والمصطلحات التي لا تحظى بالانتشار الواسع على الانترنت.

من جهة أخرى، لدينا الأداء الأسلوبي الفريد الذي تقدمه برمجيات الترجمة الآلية العصبية NMT (Neural Machine Translation)، من خلال استغلالها للترجمات البشرية الواردة في المدونات المتوازية في تقديم نماذج ترجمية بليغة للسياقات المماثلة لها عند الترجمة. أي أنها تستخرج الترجمات البشرية في المدونة المتوازية خلال العملية التوليدية للسياق الجديد المتولد من ترجمة النص المدرج (Popovic, 2017, p. 2).

في هذه الدراسة، نقترح دمج أليات الترجمة الألية السياقية مع برمجيات الترجمة الألية العصبية، قصد الاستفادة مما تقدمة الاولى من حلول لاستيعاب المصطلح وإعادة تدويره، ومما تقدمة الثانية من أسلوبية تُرجمية بليغة للخروج بنص عربي سياقُه مطابق للسياق الأصل مع المصطلحات والدلالات المكافئة في اللغة الهدف.

1.4 استثمار النهج السياقي في تطوير حوسبة صوتيات اللغة العربية:

تنظر هذه الدراسة إلى النهج السياقي بصفته أداة محتملة من شأنها الاسهام في تطوير المعالجة الآليّة للغة العربية على المستوى الصوتي، فلطالما شكلت صوتيات اللغة العربية مشكلات جمة تنبثق من كون النظام الصوتي للغة العربية فريدا ومتفردا مقارنة بالأنظمة الصوتية للغات العالم أجمع. فاللغة العربية تعد اللغة الوحيدة في العالم التي تستخدم الأحبال الصوتية جميعها في المدرج الهوائي للإنسان، حيث توزعت الأصوات بشكل عادل عليه. لكن المشكلة ليست هنا بالتحديد، بل تكمن في كون اللغة العربية تفرق كل صوت عن الأخر بدلالة معينة وبمستوى بعينه، فتتشابه طرق النطق والمعاني العامة في بعض الكلمات لكنها تختلف نوعيا في درجة المعنى المقصود في المستوى اللغوي الذي تندرج فيه. مثال ذلك ألفاظ: قَضَمَ وهَضَمَ وخَضَمَ، جميعها تحمل معناً عاما واحدا وهو أكل شيء ما، لكنها تفرق بين الطريقة التي يؤكل بها، فالقضم يعني كسر الشيء بالأسنان قبل أكله، والهضم يعني تحليل الشيء عبر لعاب الفم قبل مروره إلى الجهاز الهضمي والخضم يعني أكل الشيء نيئاً، لكن الألفاظ الثلاث تخرج من ثلاث أحبال مختلفة في الجهاز الصوتي للإنسان، فهي بذلك تساقبت في الصوت وبالتالي تساقب معناها. ففي اللغة العربية، تساقب الأصوات يولد تساقبا في المعنى بالضرورة وهو ما يضفي خصوصية بالغة لمعالجتها آليّا. وتظهر أهمية هذه الفروق أكثر ما تظهر في النصوص المتخصصة العلمية والتقنية، والتي تتطلب معالجتها دقة بالغة لأجل الحفاظ على سياقها المتخصص دون تعميم معانيها قصد تحقيق الهداف منها.

وتقترح هذه الدراسة النهج السياقي كأحد الحلول الممكنة في تمكين الحاسوب من استيعاب هذه الفروق الصوتية وخصوصيات المعاني التي تصاحبها. وقد يتأتى ذلك من خلال صف السياقات في المدونة المتوازية العربية – الأجنبية والتي كَتب نصوصها العربية كُتاب عرب، فتُترجم على اثر ذلك إلى اللغات الأجنبية وتتوازى في البرمجيات الحاسوبية لبرامج

الترجمة الآليّة ما يسهم في تعريف الحاسوب بتلك الفروق الدلالية ويسهل عليه ربطها بالسياق الترجمي المدرج، ما يمكنه في الأخير من معالجتها وتخزينها ومن ثم استخدامها في السياقات المماثلة في العمليات الترجمية ذات التوجه المشابه. من خلال هذه العملية، سيتمكن الحاسوب من تكوين قاموس برمجي أكثر دقة فيما يتعلق بتعرفه على الصوتيات العربية والفروق الدلالية التي تميزها، ما يسهم بذلك في مساعدته على إضفاء التخصصية اللازمة عند توليده للنصوص المتخصصة العربية على وجه الخصوص.

5. خاتمة:

يتضح من خلال الدراسة بأن المدونات المتوازية تشكل أداة فاعلة في إطار المساعي نحو تطوير حوسبة اللغة العربية. فهذا النوع من المدونات يوفر كما هائلا من السياقات اللغوية المتعددة والتي تشكل بدورها من خلال ذلك التوازي سياقات تُرجمية لا حصر لها، من شانها تعريف الحاسوب بالدلالات العربية في السياقات المختلفة، والكيفية التي تعالج بها اللغة العربية الدلالات العامة والمتخصصة وطرق تغليفها لما تحمله من شحنات في المصطلح العربي. إذ يمكن من خلال ما أسمته الدراسة بالنهج السياقي، الارتقاء بوحدة معالجة اللغة العربية آليّا إلى السياق مباشرة عبر الاستثمار في السياقات الترجمية المتولدة عن فعل التوازي للمدونات الأجنبية–العربية والعربية–الأجنبية على حد سواء. لذا ومن أجل تحقيق هذه المآرب، لابد من تزويد البرمجيات الحاسوبية ما أمكن بالمدونات المتوازية بمختلف انتماءاتها الميدانية، عامة كانت أو متخصصة. فبالقدر الذي يتعرف فيه الحاسوب أكثر على السياقات العربية وما تتضمنه من دلالات، بالقدر الذي أمكنه توليدها في سياقات مماثلة عبر الترجمة الآليّة تحديدا. وعليه، خلُصت هذه الدراسة إلى ما يلي من نتائج وتوصيات:

- أن النهج المعتمد حاليا في حوسبة اللغة العربية يُظهر قصورا واضحة في مواكبة ما وصلت إليه الحوسبة العالمية من تطور، مقارنة بباقي اللغات السباقة في المجال.
- ضرورة اعتماد نهج جديد يضمن للغة العربية مواكبة فاعلة للتطور الحاسوبي الحاصل.
- ضرورة الارتقاء بالوحدة المعتمدة في معالجة اللغة العربية آليّا.
- أن النهج السياقي المقترح في هذه الدراسة يمثل ألية فاعلة في سبيل الارتقاء بالوحدة في المعالجة الآليّة للغة العربية.
- للمدونات المتوازية خصائص سياقية فريدة من شأنها الاسهام في حوسبة عربية أكثر تطورا.
- أن النهج السياقي يمثل أحد الوسائل الممكنة وليس الوسيلة الوحيدة في إطار المسعى نحو تطوير حوسبة اللغة العربية، وأن تعدد الوسائل يفضي إلى نتائج أجود.

6. المراجع:

رشوان، محسن، والسعيد، المعتز بالله. (2019). *تطبيقات أساسية في المعالجة الآليّة للغة العربية*. المملكة العربية السعودية: مركز الملك عبد لله بن عبد العزيز الدولي لخدمة اللغة العربية.

علي، نبيل. (1988). *اللغة العربية والحاسوب*. الكويت: تعريب.

العناتي، وليد. (2007). *دليل الباحث إلى اللسانيات الحاسوبية العربية*. الأردن: دار جرير للنشر والتوزيع.

الغامدي، منصور بن محمد. (2017). *مدخل إلى اللسانيات الحاسوبية* (تحرير: عبد الله بن يحي الفيفي). المملكة العربية السعودية: مركز الملك عبد لله بن عبد العزيز الدولي لخدمة اللغة العربية.

Ashengo Y. A., Aga R. T., & Abebe S. L. (2021). context based machine translation with recurrent neural network for English-Amharic translation. *Machine Translation*, *35*, 19-36.

Baker, M. (1995). Corpora in translation studies: An overview and some suggestions for future research. *Target, 7*(2), 223-243.

Baldwin & Kordoni. (2011). The Interaction between Linguistics and Computational Linguistics. *LiLT, 6*(1), 1-6.

Computational Linguistics. (2021, May 17). *Britannica*. https://www.britannica.com/science/computational-linguistics

Carbonell J., Klein S., Miller D., Steinbaum M., Grassiany T., & Frey J. (2006, August 12). *Context-based machine translation* (The seventh conference of the association for machine translation in the Americas). Cambdrige. USA.

Doval, I., & Sanchez, T. (2019). *Parallel corpora for contrastive and translation studies*. Amsterdam: John Benjamins Publishing.

Leech, G. N. (2010). Corpus linguistics. *The Routledge Linguistics Encyclopedia, 3*, 103-113.

Melbi, A. K., & Foster, K. (2010). Context in Translation: Definition, Access and Teamwork. *The International Journal for Translation and Interpreting Research, 2*(2), 1-15.

Popovic, M. (2017). Comparing language related issues for NMT and PBMT between German and English. *The Prague Bulletin of Mathematical Linguistics*. 108. 209-220.

Resnik, Ph., & Smith, N. A. (2003). The web as a parallel corpus. *Computational Linguistics*, *29* (3), 349-380.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Zanettin, F. (2011). Translation and corpus design. *SYNAPS, 26*, 14-23.