الترجمة الرقمية: انعطافة عضوية في مسيرة الترجمة الآلية

أ. د. حسام الخطيب

1. تمهيد

لوحظ في السنوات الأخيرة من القرن العشرين وفي مستهل القرن الحالي التفات قوي في الوطن العربي باتجاه الترجمة. وبعد أن كانت الترجمة مهملة نسبياً في مجال النشاط الثقافي العربي، كما تدل على ذلك المتابعات المسحية لليونسكو، ولاسيما الإصدار الدوري: فهرس الترجمة **Index translationum**، الذي يتابع تطورات نشر الكتب المترجمة في جميع أنحاء العالم.

وكنت بيّنت في دراسات سابقة لي أن القوائم المنشورة في هذا الدليل حول النشاط الترجمي في البلاد العربية تشير إلى فقر شديد في هذا المجال بالمقارنة مع بلدان العالم الأخرى من الناحيتين الكمية والكيفية، ولا يوجد أثر لخطط مستقبلية أو تخصصية، اللهم إلا ما تفعله بعض وزارات الثقافة والهيئات الرسمية أو شبه الرسمية من ناحية تخصيص سلاسل معينة حسب موضوعات مفصلة مع تفاوت كبير بهذا الشأن بين كل قطر عربي وآخر. ويأتي في مقدمة هذه الخطط المشروع المتميز للمجلس الأعلى للثقافة في مصر العربية. إلا أن مثل هذا المشروع وسواه من مشروعات أخرى أضيق نطاقاً مثل الجهود الترجمية لوزارات الثقافة في كل من سوريا والعراق قد مضت دون أي تقييم منهجي سوى التقييم الكمي والمعرفي أي عدد المنشورات المترجمة وأحياناً تنوعات تخصصها.

وهناك إجماع واضح على أن الترجمة كانت (وربما ما زالت) شبه سائبة، أ ي يختلط فيها الحابل والنابل، وتفتقر إلى معايير للتقييم، سواء من ناحية الأمانة العلمية أو من ناحية السلامة اللغوية للغة الهدف **Target Language** ، وهي غالباً اللغة العربية.

وهناك استثناء وحيد هو ما تنشره المجلات الأدبية بين حين وآخر من مساجلات ومشاحنات بشأن بعض الترجمات – ولاسيما الأدبية والفكرية – التي ينشرها أساتذة ومترجمون لهم مكانة رفيعة. وبوجه عام ظلت النظرة إلى الترجمة مشوبة بشيء من التصنيفية السلبية، أي اعتبار الترجمة عملاً مكملاً وبعيداً عن الإبداع ولا يستحق التدقيق.

وحتى يوم الناس هذا يصعب أن يطمئن المرء إلى وجود تغيُّر في الموقف العام من الترجمة، اللهم إلا من الناحية الكمية التي تتمثل في الجوانب التالية:

- زيادة ملحوظة في عدد المقبلين على مهنة الترجمة بينما كانت تعدّ مهنة إضافية، أي عملاً مكملاً لأنواع أخرى من المهن ولاسيما تدريس اللغات الأجنبية.
- زيادة في كمية الإنتاج الترجمي، أي في عدد الترجمات المنشورة، إلا أن هذه الزيادة يجب أن تقاس من خلال منظور نسبي أي ليس بالمقارنة مع الوضع السابق في الوطن العربي، ولكنها إذا قورنت بالزيادة شبه الخيالية في النشاط الترجمي على المستوى العالمي تبدو باهتة شأنها شأن أنواع النشاط المعرفي الأخرى في الوطن العربي.
- قفزة في مجال مؤتمرات الترجمة على المستوى العربي الشامل في بضع السنوات الأولى من القرن الجديد، على حين كانت هناك ملتقيات محدودة في بعض الأقطار العربية، أو كانت هناك بنود محددة حول الترجمة كما حدث أحياناً في: مؤتمرات الاتحاد العام للأدباء والكتاب العرب. وذلك كله إلى جانب نشاطات محدودة في بعض الجامعات العربية. ويلاحظ أنه خلال السنوات القليلة الماضية عقدت مؤتمرات مهمة جداً في مجال الترجمة على المستوى العربي الشامل. وذلك إلى جانب نشوء مؤسسات متخصصة بالترجمة، غالباً بفضل أفراد أثرياء محبين للمعرفة، وأحياناً بمبادرة من هيئات المجتمع المدني بدعم من العاملين في مجالات الترجمة، وأحياناً من خلال مؤسسات عربية جامعة مثل المنظمة العربية للتربية والثقافة والعلوم.
- وهكذا جاءت القفزة المذهلة من خلال تعدد مراكز الترجمة ومؤسساتها في الوطن العربي، كأنما العرب أفاقوا فجأة على أهمية الترجمة، وبدأوا كالعادة يتنافسون في افتتاح المراكز، وإنشاء

دبلومات الترجمة في الجامعات[1] ولكن دون أي اهتمام بالتنسيق أو التكامل أو تجنب تكرار الجهود والتمويل للهدف الواحد.

هذا من الناحية الكمية، وكله خير وبركة. ولكن يلاحظ أن قانون التحولات الكمية باتجاه تبلور كيفي يصعب أن ينطبق على الإنتاج الترجمي الحالي في البلاد العربية، حتى في مجال الترجمات الأدبية والفكرية ذات الموضوع الراقي، وذلك بسبب فقدان المعايير وغياب أي منظومة للتقييم والنقد متفق عليها اتفاقاً ملفوظاً أو ملحوظاً. ونادراً ما نقرأ في الصحافة الأدبية أية ملاحظات حول المستوى الفني واللغوي للترجمات، بل مستوى الأداء والأمانة، باستثناء المشاحنات التي أشرت إليها سابقاً. ومن المأمول أن يقود التقدم الملحوظ في الدراسات النظرية حول الترجمة إلى التوصل لما يشبه المعايير العامة.

أما في الترجمة الرقمية التي هي موضوع هذه الورقة فالملاحظ أنه مع اتساع تجارة (بزنس) الترجمة الآلية على الانترنت، بدأت محاولات وضع أنساق للتقييم مستقاة من طبيعة عمل الترجمة الآلية التي تجمع بين صرامة الآلة وحرفيتها من جهة وغياب الحساسية اللغوية والجماليات التعبيرية من جهة أخرى. ومن هنا كانت أهمية استحداث معايير للتقييم متجانسة مع الطبيعة المزدوجة للترجمة الآلية التي ينتظر أن يكون لها في المستقبل دور قوي في التفاعل والتواصل بين البشر بوجه عام، وكذلك فيما بين تطورات الأنظمة المعرفية، وتبادلات الخبرة على المستوى العالمي. ومن هنا كان اهتمام الورقة الحالية بإعطاء فكرة ولو بسيطة عن تطور معايير التقييم ومناهجه في البلدان المتقدمة. ومن الواضح أن هذا البعد المهم لم يلق حتى الآن أي نوع من الاهتمام المؤسسي في الأقطار العربية. مع العلم أن التقييم في البلدان المتقدمة لقي في السنوات الأخيرة الماضية اهتماماً متصاعداً من جهات محايدة ومعترف بها، كالمعاهد الرسمية التي سنشير إليها في الورقة الحالية. مع التأكيد أن الشركات

[1] . تذكرنا هذه المنافسة والتعددية بمقالة ميخائيل نعيمة المشهورة في أوائل القرن الماضي بعنوان "فلنترجم". إذ تأخرت قراءتها حوالي قرن من الزمان. كما أود أن أشير إلى أن هذه الظاهرة التنافسية التكرارية واردة جداً في مختلف مستويات العمل السياسي والاجتماعي والثقافي العربي. وقد أسميتها في بعض دراساتي بظاهرة التفكير الموجي عند العرب.

التجارية تقوم بهذا العمل منذ زمن من خلال النشاط التنافسي الذي لا يرحم بين مئات الشركات العاملة في مجال الترجمة الآلية ولاسيما على الشابكة الدولية.

ومن خلال مسعى متواضع لمتابعة التطورات المتسارعة في مجال الترجمة الآلية، تقدم هذه الورقة عرض لنقلة ذات أهمية جوهرية في تطور الترجمة الآلية، وهي الاعتماد على الترجمة الرقمية والإحصائية على أمل تجنُّب صعوبات الالتواءات اللسانية واللغوية التي تعجّ بها لغات العالم وتعترض مسيرة الترجمة الآلية التقليدية، مع العلم أنه برزت مجدداً محاولات للجمع والتوازن بين هذين المجالين اللغوي والرقمي، وإن كانت مازالت في مهدها مع بروز مشروعات تطورات متلاحقة في مجال الترجمة.

2. هل تكون الترجمة الرقمية هي الحل

لم تكن الآلة أبداً هي العقبة في وجه تطور الترجمة الآلية بالسرعة التي يتطلبها عصرنا اللاهث دون هوادة. ولكن كان قصور الدراسات اللسانية التي توفر المعالجة الآلية للغة هو المشكلة دائماً.

وفيما يتعلق باللغة العربية تتضخم هذه المشكلة إلى درجة مأزقية بسبب فقر الدراسات اللسانية المتعلقة بالعربية [2] ويبدو أن مغامرات التنافس التجاري بين الشركات العملاقة في بلاد الشمال، وفي الولايات المتحدة بوجه خاص، إلى جانب قوة المغامرة الابتكارية

[2] . انظر الملحق (رقم – 1) الذي يبين احتياجات اللغة العربية لدراسات اللسانيات الحاسوبية. والملاحظ أن بعض شركات الترجمة الآلية بدأت بتلبية هذه الاحتياجات، ولكنها تعثرت وتخلت عن المشروع وبقيت جهود بعض الباحثين الذين عملوا في هذا المجال تنتظر الإفادة من مبادراتهم من خلال توحيد الجهود العربية المبعثرة . وسوف يتضح هذا الأمر من خلال البيانات اللاحقة (ملحق – 3) .

دفعت بعض رُوّاد الابتكار مثل كيفن نايت Kevin Knight ودانييل ماركو Daniel Marcu وشركاءهما باتجاه تجريب الحل الرقمي للقفز وراء العقبات. وقد بدأت المغامرة بأستاذين جامعيين أقدما على إنشاء مشروع شركة لصناعة برمجيات ترجمة آلية معتمدة على الإحصاء الرقمي من شأنها أن تترجم النصوص من لغة إلى أخرى. وحين عرض الأمر على المستثمر برايس بنجامين Bryce Benjamin في الشهر الأخير من عام 2001 أدرك بسرعة أهمية هذا الاختراع، ولاسيما من ناحية تلبيته لمتطلبات وزارة الأمن الداخلي الأمريكية (وهي وليدة أحداث الحادي عشر من أيلول) وكذلك من ناحية تزايد الحاجات التجارية للاتصالات بين الشركات الكبرى في العالم وتوسع سوق التجارة العالمية بتأثير تسهيلات الإنترنت.

واليوم أصبحت شركة " الحائك اللغوي Language Weaver" الثلاثية التي بدأت بخمسة وثلاثين فنياً، واحدة من الشركات القيادية البارزة في مجال الترجمة الآلية المزدهرة بتسارع مذهل. وهناك تنافس حاد في مضمار سرعة تلبية احتياجات الترجمة لصالح حملة الولايات المتحدة على الإرهاب، حيث يعجز المترجمون البشريون عن القيام بالمهمات الصعبة لترجمة ذلك الفيض الهائل من أشرطة التسجيل والبيانات والوثائق، التي يقاس مقدارها بالملايين. وهكذا أصبحت البرمجيات الآلية هي الاختيار الأفضل للمتنافسين. وأقدمت شركات عملاقة متعددة الجنسيات على ارتياد المغامرة البرمجية، وتصدرت شركات مثل غوغل وياهو وسيستران وآي بي إم IBM قائمة المتنافسين على المستوى العالمي في عملية تلبية الطلب المتزايد في مجال ترجمة الوثائق التقنية وصفحات الإنترنت.

إن برمجيات " الحائك اللغوي" تترجم من وإلى الانكليزية في مجال عدة لغات. وبالطبع ولأسباب سياسية وعسكرية ونفطية تتصدر اللغة العربية هذه القائمة وتليها الصينية والإسبانية. وفي المرحلة الأولى اقتصرت مهمة متابعة البرمجيات الرقمية على تحديد أهمية الأخبار التي تبثها وسائل الإعلام باللغات الأجنبية وتمييز ما هو ضروري من

خلال تقديم فكرة عن مضمونها. وهي مرحلة ابتدائية طبعاً، أخذت تتسارع تدريجياً حتى احتلت قصب السبق في حلبة التنافس المحموم لتلبية احتياجات عجلة الأمن القومي الأمريكي.[3]

[3] . انظر الملحق – 2 الذي يبين تسلسل كفاءة الترجمة من العربية إلى الانكليزية مع تطور البرامج سنوياً .

3. الترجمة الرقمية

تمتاز الترجمة الرقمية بالبساطة، من المعروف أن معظم أنظمة الترجمة الآلية ظلّت حتى مطلع القرن الحالي تعمل من خلال الكلمات المفردة أو تستخدم مجموعات معقدة يقوم اللسانيون بإعدادها وترميزها دلالياً وتركيبياً ونحوياً بغرض تلقيمها للبرمجيات وهي عملية معقدة جداً ومكلفة وبطيئة التطور. أما تقنية " الحائك اللغوي" فهي تستخدم الترجمة البشرية أساساً لعملية الترجمة الآلية معتمدةً على وثائق هيئة الأمم المتحدة المترجمة، وذلك باختيار مقاطع من لغتين في موضوع واحد وتنسيقها في اصطفاف كل عبارة phrase مقابل عبارة نظيرة من خلال ترميز رقمي N-gram. واستناداً إلى عملية المقارنة هذه تتعلم البرمجيات كيف تنتقل من أنموذج لغوي إلى نظيره في اللغة المستهدفة[4]. فمثلاً إذا كانت هناك مجموعة كلمات محددة تؤدي معنى معيناً بالعربية فإن البرمجية تبحث عن صف الكلمات المناظرة في اللغة المستهدفة وتختار ما يناسب لغة المصدر المرسلة. وهذا يعني أن النظام المعتمد يترجم العبارات ضمن سياقها في اللغة المرسلة وليس المفردات من لغة إلى أخرى، فمقابل عبارة interest rate مثلاً لا نجد الاحتمالات المعجمية الأصلية الخاصة بإحدى الكلمتين مثل الاهتمام والسرعة curiosity and speed بل نجد العبارة الصحيحة وهي سعر الفائدة. والقاعدة لإغناء النظام هي أن تضيف إليه باستمرار التغذية اللغوية من العبارات الكافية للأداء المقصود، ويتم ذلك بالتدريج انطلاقاً من تغذية أساسية بسيطة إلى تطورات أوسع. وتعد وحدة N-gram الوحدة الأساسية التي تشترك فيها الانظمة المختلفة للترجمة الرقمية. وهي نوع من العبارات عدداً محدداً من الكلمات التي يجري ترتيبها وفق أساس معين من خلال العلاقة النوعية فيما بينها. ومن هنا أتت التسمية الرقمية أو الإحصائية، وهكذا تقاس كفاءة الترجمة الرقمية بعدد وحدات الكلمات N-gram التي

---

[4] . انسجاماً مع المصطلحات العالمية الشائعة نستعمل كلمة لغة مرسلة مقابل source language ومستهدفة مقابل target language .

يتعامل معها النظام والتي تعتمد أصلاً على ترجمة بشرية ذات مستوى معترف به، وذلك بعد إعدادها بترتيب نوعي يناسب الهدف المقصود. ومرة أخرى يعد هذا النظام أبسط ما عرف في الترجمة الآلية من أنظمة.

وحتى الآن مازال هذا النظام الإحصائي يتطلع إلى مضاعفة سرعة الأداء ليتناسب مع الحاجات العصرية المتزايدة. ولكن إذا كان برنامج " الحائك اللغوي" يستطيع أن يترجم من خلال حاسوب شخصي خمسة آلاف كلمة في الدقيقة على مستوى من الدقة المتحسنة باستمرار فإن شبكة من المخدِّمات servers تستطيع أن تترجم على النحو نفسه 500 ألف كلمة في الدقيقة الواحدة.

وبفضل هذا النظام لن يفقد المترجمون البشريون وظائفهم لأن " الحائك اللغوي" يعتمد أصلاً على الترجمة البشرية ذات المستوى المتفق عليه، ولكن بالطبع تشتد الحاجة إلى مترجمين في أعلى مستوى من الإتقان لتكون ترجماتهم أساساً للمطابقة التي يقوم بها الحائك اللغوي ويختزلها وينطلق منها إلى لغات أخرى أي أنه مقيد بطبيعة النصوص التي يتغذى بها. وينبغي الاعتراف أن هناك شكوكاً كثيرة تدور حول دقة مثل هذه الترجمات. ولكن التغذية الراجعة من قبل المؤسسات الأمنية التي جرى التعامل معها أتت مشجعة جداً، بل ساعدت على التقاط معلومات ذات أهمية خاصة وفي وقتها المناسب. وتحيط الوكالات الأمنية الأمريكية أسماء الشركات التي تتعامل معها في هذا المجال بسياج من السرية.

4. تاريخ قريب وتطورات سريعة

وعلى أية حال بدأت الاستفادة الفعلية من الحائك اللغوي عام 2003. وكذلك أخذت تتطور عنه عدة آلات جديدة مثل المترجم الفوري النقال الذي يستخدمه الأطباء والمرضى في حالة اختلاف اللغة (ولاسيما الانكليزية والاسبانية في أمريكا) إلى جانب أدوات عملية ناطقة في مجالات مختلفة مثل مجال الترجمة من الكلام المنطوق المذاع إلى وثيقة تحريرية. وقد تمّ إحراز نجاح فائق في هذا المجال .

وتجري الآن جهود لتحويل هذا الجهاز من الثنائية اللغوية إلى تعددية اللغات. ولا بد من أن نتذكر هنا وجود اتفاق عام على أن ما تنتجه الترجمة الآلية بكل أشكالها لا يقدم لغة سليمة. وهنا تصدّى صاحبا النظرية الرقمية كيفن نايت وماركو لمعالجة هذا الأمر، وجرى تخصيص 285000 دولار، لترويض الآلة المترجمة على مراعاة أصول النحو في مجال اللغة الانجليزية (المعالجة النحوية Grammatical Processing). وفي تقديرهم أنه يمكن إنجاز هذا العمل وتقديمه قبل نهاية عام 2006 إلى المعهد القومي للعلم والتكنولوجيا[5] NIST من أجل تقييمه[6].

ومن الضروري التذكير هنا أن التجارب الترجمية الناجحة من خلال الحائك اللغوي وغيره من التقنيات تتطور باستمرار وتتعرض لتقييم مؤسسي دقيق. ويبدو نصيب اللغة العربية فيها حسناً من الناحيتين الكمية والكيفية، إلا أن التقدم الملموس يكاد يكون محصوراً بالعربية حين تكون لغة المصدر المرسلة وليس لغة الهدف أي في مجال الترجمة من العربية إلى الانكليزية مثلاً وليس العكس. وهذه النقطة لا تخدم العربية ولا

---

[5] . بدأ العمل بهذا المعهد عام 2002 ، ويتألف من مجموعة مؤسسات رسمية وجامعية وتجارية تصدر كل سنة في الشهر الثامن تقييماً مفصلاً لتطورات الترجمة من مختلف الجوانب، ويتمتع المعهد باحترام واسع، وبذلك سبقت الترجمة الآلية الترجمة البشرية في مجال التقييم.

[6] . بتصرّف عن مجلة Technology Review.com

قضية ترجمة العلوم مثلاً إلى العربية لأن الوضع الحالي للإنتاج العلمي العربي هو وضع المستورد حتماً وليس المصدّر. والملاحظ أن معظم اللغات الآسيوية تعاني من المشكلة ذاتها بسبب اختلاف تركيبها عن اللغات الأوروبية، ويلاحظ بوجه خاص أن اللغتين اليابانية والصينية تعانيان أيضاً، وربما بصورة مضاعفة، تتناسب مع طبيعتهما النحوية المعقدة وضرورة اعتمادهما على الترجمة الآلية، وهذا ما يؤخر استخدام الترجمة الآلية بوجه عام في المجالات العلمية والمجالات التي تحتاج إلى دقة فكرية.

وهنا تنبغي الإشارة إلى أن التقنية الرقمية غير محصورة بالترجمة الآلية من العربية إلى الإنكليزية، بل أخذت تتوسع بالتدريج لتشمل مروحة واسعة من اللغات التي لم تستطع الترجمة الآلية سابقاً ترويضها، ومن أبرزها اللغات الآسيوية في دار الإسلام كالأوردية ولغات مناطق باكستان وأفغانستان المختلفة، وكذلك الصينية واليابانية والكورية. وحتى الآن عقدت في الصين بالذات[7] عدة مؤتمرات لهذا الغرض وحددت هدفها بكسر الحاجز اللغوي وتعزيز التفاعل بين ثقافات العالم المعاصر. ومثل ذلك حدث في أمريكا الجنوبية.[8]

[7] . انظر بوجه خاص المجلة المتخصصة الصينية Practices for Digital Asia, Machine Translation Mediated.

[8] . وذلك من خلال مؤتمر " الترجمة الآلية في الأمريكيتين، MT in the Americas, Aug. 2006 ".

5. التقنية الرقمية طريق المستقبل

بقي التذكير بأن عالم المستقبل هو عالم التقنية الرقمية بوجه عام وبذلك تكون الترجمة الرقمية أداة من أدوات بناء مستقبل عالم بلا حدود.

ومن الأمثلة الرقمية التي تتصل اتصالاً مباشراً بالترجمة الآلية وترجمة العلوم صدور مجلة "العلوم الأمريكية American Scientific " ، بعدة لغات من خلال الترجمة الآلية، وحتى الآن تتراوح الدقّة في هذه الترجمات بين 70–80 %، وهي مجلة مواكبة علمية ثمينة، وتصدر طبعتها العربية عن المجلس الوطني للثقافة في دولة الكويت ويسهم في ترجمتها علماء عرب مشهود لهم بالدفة والدراية.[9]

إن الاتجاه الرقمي في الترجمة الآلية يتصاعد يومياً في الولايات المتحدة ونسبياً في بريطانيا، ويشكل ثورة معرفية تقانية كاسحة. وغالباً ما تقود هذه الثورة مراكز الأبحاث في الجامعات الأمريكية، وكذلك مراكز البحوث المدنية والعسكرية والاستخباراتية. وهناك وكالة أمريكية متخصصة بالإشراف على جميع البحوث التي لها علاقة بالأمن القومي وما يسمى بالحرب على الإرهاب، وهي "وكالة مشروع البحث المتعلق بالدفاع The Defense Advanced Research Project Agency (DARPA) ". ونتفق هذه الوكالة مبالغ طائلة على البحوث المتعلقة بالمصالح العليا للولايات المتحدة، وتدخل في هذا النطاق الترجمة إلى الانكليزية من العربية والأوردية (لغة باكستان) واللغات الأخرى في منطقة أفغانستان وما حولها، كما ذكرنا سابقاً.

وفي بعض الأحيان تتولى مراكز البحوث في الجامعات عملية التنفيذ بعد التوصل إلى نتائج مشجعة، ولكن في أكثر الأحيان، وحين تصبح نتائج المختبرات مبشرة بوعود

---

[9] . أغلب هؤلاء من قدامى متخرجي الجامعات السورية، وقد درسوا العلوم أصلاً باللغة العربية، ثم أكملوا دراساتهم العليا في بلدان أجنبية.

تجارية رابحة تقدم الشركات الكبرى على تبني هذه المشروعات، وفي مقدمتها بالنسبة للغة العربية: سيستران وغوغل وأي بي إم IBM وسواها.

وهذا يعني أن الترجمة الآلية الرقمية هي جزء من مشروع الثورة الرقمية الكبرى في عالم التقانة والمعرفة. ومن أبرز الأمثلة على ذلك المشروعان الكبيران اللذان أنجزتهما جامعتا ييل وكورنل، وسوف نعرض لهما باختصار رغبةً في وضع مشروعات الترجمة الرقمية في إطارها العام.

أ. مشروع جامعة ييل Yale :

أصبح هذا المشروع جاهزاً وتمكن متابعته على الانترنت من خلال موقع مكتبة الجامعة، ويتضمن أرشيفاً شاملاً ومفصلاً للإنتاج التحريري والإذاعي والتلفازي على امتداد العالم كله، وذلك من خلال نظام DVB (Digital Video Broadcasting). ويخدم هذا المشروع أغراضاً علمية معرفية وكذلك أغراضاً أمنية وتجارية واقتصادية، وذلك من خلال متابعة تقارير الأسواق والجهات الصناعية والتجارية على امتداد العالم كله. ولكل قارة أو منطقة جغرافية سجلها الخاص بأقاليمها المختلفة. ومنها الشرق الأوسط والمناطق العربية والإسلامية الأخرى التي أصبحت مركز اهتمام عالمي لأسباب سياسية وعسكرية. إلا أنه لا دخل لهذه الأسباب بصدقية المعلومات التي يقدمها البرنامج من خلال المعطيات الموضوعية. وهنا يمكن القول إن العالم بأجمعه مخزن رقمياً في مكتبة جامعة ييل.

ب. مشروع جامعة كورنيل Cornell الرقمي

وهو مشروع رقمي جيوفيزيائي وجيولوجي، منسق رقمياً، ويقدم خرائط ومعلومات ومراجع وافية عن الشرق الأوسط بلغات مختلفة، ومن الملاحظ أن جزءاً كبيراً من المراجع هو أصلاً رسائل جامعية أو بحوث أجريت في الغرب (أمريكا وبريطانيا بوجه خاص)، على يد باحثين ينتمون إلى المناطق التي تتناولها البحوث غالباً.

ويلفت النظر في المشروع وجود بحث مستقبلي حول جيولوجيا الشرق الأوسط وعلاقته بالبترول من خلال متابعات ميدانية حيّة.

وبالمقارنة بين مشروعات التحول الرقمي في العالم والمشروعين المقدمين أعلاه، نجد أن الترجمة الرقمية تمثل مواكبة لغوية للتطورات العالمية الراهنة. ويجب أن تؤخذ تطوراتها بجدية كاملة، وسيكون لها في المستقبل تأثير كبير في مشروعات ترجمة العلوم. ومن المعروف أن الجامعات الأمريكية بدأت تتنافس في الانتقال إلى التقنية الرقمية التي جعلت التوصل إلى محتويات مكتباتها متاحة للمستفيدين من الإنترنت. وهذه الظاهرة أشبه بثورة معرفية شاملة بوصفها تكسر الاحتكار العلمي الذي كان حتى الآن سائداً في العالم.

6. الحضور البشري في الترجمة الرقمية أولاً وآخراً

ربما كان التحول إلى الترجمة الرقمية يحدد نقلة نوعية مختلفة عن الاتجاه الأول للترجمة الآلية الذي بقي طوال نصف القرن الماضي يعمل بجد واجتهاد على التوصل إلى نمط من الترجمة الآلية مستقل في أدائه عن العنصر البشري استقلالاً شبه تام. وقد أتت الترجمة الرقمية بعد تعثر كل تلك الجهود السابقة عن التوصل التام إلى هدفها، أتت لتجعل بداية العملية هي الانطلاق من الترجمة البشرية الجيدة والمتفق على صحتها لتكون أنموذجاً يحتذى في الترجمة الآلية للنصوص على أساس المشابهة. وأضيف مجدداً إلى هذا العنصر البشري عنصر ختاميّ يقوم على مبدأ الاستفادة من الجهود البشرية في الترجمة وجعلها المرجع النهائي في تحديد صلاحية الترجمة الآلية. وقد توصلت مجموعة من الباحثين إلى طريقة جديدة تعتمد في مجال الترجمات الأدبية والفكرية على الاستفادة من مجمل الجهود البشرية السابقة في هذا المجال، وقطف ثمرتها من خلال جعل المتوسط المشترك لهذه الجهود هو الأصل الذي ترجع إليه الترجمة الآلية الرقمية. فمثلاً جرى نقاش حول ترجمة رواية دوستويفسكي "الجريمة

والعقاب" التي تفاوتت بشأنها اجتهادات المترجمين البشريين باللغة الانكليزية. وتمّ تلقيم الآلة بمختلف هذه الترجمات وطلب إليها أن تختار أكثر العبارات تردداً في الترجمات المختلفة وأن تجعلها المرجعية النهائية. وبذلك تكون الترجمة الرقمية هي صاحبة الفضل في إعادة الاعتبار للترجمة البشرية وكذلك في تضييق الهوة بين حصسلتَيْ الترجمة البشرية والترجمة الآلية.

ولكن هذه اللعبة تثير إشكالات كثيرة. ومن أهمها أن الترجمات البشرية الأدبية والفكرية تتفاوت عادة في مستواها تفاوتاً شديداً وبعضها يرتفع إلى مستوى القمة وبعضها يتعثر تعثراً شديداً. والمساواة بينها في الاختيار مسألة ديمقراطية ولكنها تنطوي على تجاوز. ونعتقد أن الأصح هو اعتماد ترجمات جرت تزكيتها إجماعاً على يد ومتخصصين وإدخالها في هذه العملية لتكون النتيجة سليمة.

وهنا نود أن نذكر أن الترجمات البشرية تجور أحياناً على النصوص، وتمضي دون معيار متفق عليه، أي أن كل ترجمة بشرية بحد ذاتها ليست أفضل من الترجمة الآلية، وقد شهدت الأسواق ترجمات متسرعة وغاية في الرداءة الاستيعابية واللغوية. وينطبق ذلك على الترجمات التجارية التي تعج بها أسواق المكتبة العربية، حيث تمضي الترجمات دون أي تحكيم أو معيار يميز الصالح والطالح.

ملحق – 1

حول المعالجة الآلية للغة العربية

نعرض فيما يلي خلاصة للجهود التونسية الراهنة في مجال المعالجة الآلية للغة العربية:

- حول العلاقات المتشابكة بين اللسانيات الحاسوبية والمعلوماتية:
  - المعالجة الآلية للكتابة.
  - المعالجة الآلية للصرف.
  - ميكنة المعجم.
  - المعالجة الآلية للنحو.
  - المعالجة الآلية للدلالة.
  - المعالجة الآلية لسياق المتصل.
- حول المعالجات الآلية للسانيات الحاسوبية:
  - معالجة البيانات (أو المعطيات).
  - معالجة الكلمات.
  - معالجة النصوص.
  - استرجاع المعلومات.
  - معالجة المعارف.
  - الترجمة الآلية.
- حول التطبيقات الثانوية نذكر على سبيل المثال:

النظم الآلية للنشر المكتبي، والتعليم بمساعدة الحاسوب.

وتتدرج البحوث التونسية في هذا الإطار من المعالجة الآلية للغة العربية. وقد شملت هذه البحوث والدراسات المعالجات التالية:

أ. المعالجة الآلية للكتابة العربية:

تناول الباحثون في هذا المجال ما تميزت به منظومة الكتابة العربية عن نظيراتها من المنظومات الأخرى مثل:

- ارتباط الكتابة العربية بالأصوات الوظيفية.
- اختلاف الأبجدية العربية عن نظيراتها في استخدامها لعلامات التشكيل.
- خاصية التشبيك وقواعد الخط اليدوي التي التزمت بها الطباعة العربية وكذلك النظم الآلية.
- توليد أنماط أشكال الحروف العربية تلقائياً والتعرف عليها آلياً.
- تحويل الكتابة العربية إلى كتابة بطريقة "برايل".

ب. المعالجة الآلية للصرف العربي:

نظراً لأهمية منظومة الصرف العربي وما تتطلبه من معالجة آلية، فقد اهتم الباحثون بمجموعة من التطبيقات، من أهمها:

- ضغط النصوص العربية بالأسلوب الصرفي.
- استرجاع النصوص العربية بالأسلوب الصرفي.
- تصحيح الأخطاء الإملائية بالأسلوب الصرفي – النحوي.

ج. ميكنة المعجم العربي:

يعتبر إعداد بنك للجذور العربية بواسطة الحاسوب من أهم الأشياء التي تحقق وحدة المعجم العربي، xوأداة العمل الضرورية لكل عمل في مجال اللسانيات بمستوياتها الصوتية والصرفية والمعجمية والدلالية،x بهدف التوصل إلى معاجم آلية للغة العربية. ويحتوي هذا البنك للجذور العربية على المعاجم التالية:

1. المعجم النظري الآلي.

2. المعجم الآلي للجذور المدونة في أمهات المعاجم العربية (الصحاح للجوهري، ولسان العرب لابن منظور، وتاج العروس للزبيدي).
3. معجم الجذور الثلاثية المهملة.
4. معجم الجذور المرفوضة.
5. معجم الاشتقاق الأكبر.

د. المولد الآلي للكلمات العربية الأساسية والمستحدثة:

انطلاقاً من بنك الجذور العربية الثلاثية، وقواعد تكوين الكلمات في اللغة العربية، تمّ إنجاز مولد آلي للكلمات العربية الأساسية بواسطة الجذور والأوزان، ومولد للكلمات المستحدثة بواسطة الكلمات الأساسية والإلصاق، أي بزيادة السوابق واللواحق والزوائد الوسطية. وقد استثمر الباحثون هذه المجموعة من القواعد التي تعتمد على عمليات التحول الداخلي والاشتقاق والتركيب في توليد كلمات مستحدثة من كلمات أساسية.

وأخيراً، فإن ما تم إنجازه من بحوث تونسية في مجال المعالجة الآلية للغة العربية التي ترتكز على أسس اللسانيات الحديثة وعلوم الحاسوب، قد اقتصر على الجوانب الرئيسية التي تهم الباحث في مجال اللسانيات الحاسوبية قصد تعريب المعلوماتية.

ولضمان نجاح هذه البحوث العلمية في مجال المعالجة الآلية للغة العربية، وللنهوض بلغتنا العربية حتى تساير التقدم العلمي والتكنولوجي المتسارع، يجب علينا معالجة بقية المنظومات الأخرى للغة العربية التي لم يتم تناولها حتى الآن بالبحث والتحليل، مثل المعالجة الآلية للنحو العربي، والمعالجة الآلية للدلالة، ومعالجة السياق العربي المتصل آلياً...[10]

[10] . خلاصة تقرير من إعداد الدكتور حسين الهبائلي . والمأمول أن تتضافر الجهود العربية في هذه المجالات ولاسيما الجهود الفائقة التي يتولاها الدكتور نبيل علي في مصر، وكذلك جهود مؤسسة صخر.

Sample Translations from Arabic > English
Using Language Weaver's SMTS for Arabic

The following tables illustrates the improvements in the Arabic translations from our first version, Version 2.4 through Version 4.0.

Original text	**واضاف المصدر ذاته ان فلسطينيا اوقف الليل الماضي في رام الله وبحوزته حقيبة تحتوي على حزام متفجرات زنتها 20 كيلوغراما.**
Human Translation	**The same source added that a Palestinian was arrested last night in Ramallah with a bag containing an explosive belt weighing 20 kilograms.**
Version 2.4	The source said, adding that Palestinians were arrested overnight in Ramallah possession portfolio contains a belt of explosives-20 kilograms.
Version 3.0	He added that Palestinians were arrested last night in Ramallah with a bag containing a belt of explosives blown apart by 20 kg.
Version 3.1	He added that Palestinians were arrested last night in Ramallah with a bag containing a belt of explosives blown apart by 20 kg.
Version 4.0	The same source added that the Palestinians were arrested last night in Ramallah in possession a bag containing the explosive belt-20 kg.

Original text	**وقال خلال مؤتمر صحافي في ختام لقاء مع وزير الخارجية الفرنسي دومينيك دو فيلبان ان المفتشين "يحتاجون الى بضعة اشهر لاتمام مهمتهم".**
Human Translation	**He said during a press conference at the end of a meeting with the French Foreign Minister, Dominique de Villepin, that the inspectors "need several months to complete their task".**
Version 2.4	He said during a press conference after talks with French Foreign Minister Dominique de Villepin said inspectors "will need a couple of months to complete their mission".
Version 3.0	He said during a press conference at the end of a meeting with French Foreign Minister Dominique de Villepin that inspectors "in need of a few months to complete their mission."
Version 3.1	He said during a press conference at the end of a meeting with French Foreign Minister Dominique de Villepin that inspectors "in need of a few months to complete their mission."
Version 4.0	He said during a press conference at the end of a meeting with French Foreign Minister Dominique de Villepin that inspectors "need a few months to complete their mission."

Original text	**وعارض ممثلو ايران علانية التهديد الامريكى بشن حرب ضد العراق قائلين انها ستدمر الاستقرار والامن فى الشرق الاوسط.**
Human Translation	**Representatives of Iran openly opposed the American threat to launch a war on Iraq, saying that it would destroy stability and security in the Middle East.**
Version 2.4	opposed the representatives of Iran publicly announced US threats of waging war against Iraq, stating that it would destroy stability and security in the Middle east.
Version 3.0	Disagreed with the representatives of Iran publicly American threatening to wage war against Iraq, saying that it would destroy stability and security in the Middle East.
Version 3.1	Disagreed with the representatives of Iran publicly American threatening to wage war against Iraq, saying that it would destroy stability and security in the Middle East.
Version 4.0	Representatives of Iran publicly opposed the American threats to wage war against Iraq, saying that it would de-stroy stability and security in the Middle East.

Original text	**بغداد 1-15 (اف ب) - اعلن متحدث باسم وزارة الخارجية العراقية ان قيام المفتشين الدوليين بزيارة حي قريب من قصر رئاسي في بغداد يشكل "خطوة استفزازية".**
Human Translation	**Baghdad 1-15 (AFP) - A spokesman for the Iraqi Foreign Ministry stated that the visit by the international inspectors to a district close to a presidential palace in Baghdad was "a provocative step".**
Version 2.4	Baghdad 1-15 (AFP)-Iraqi Foreign Ministry spokesman spokesman said that a UN arms inspectors to visit the neighborhood near the presidential palace in Baghdad poses "a provocation".
Version 3.0	Baghdad 15-1 (AFP)-A spokesman for the Iraqi foreign ministry that the international inspectors to visit district near the presidential palace in Baghdad was a "step provocative".
Version 3.1	Baghdad 15-1 (AFP)-A spokesman for the Iraqi foreign ministry that the international inspectors to visit district near the presidential palace in Baghdad was a "step provocative".
Version 4.0	Baghdad 1-15 (AFP)-- announced a spokesman for the Iraqi Foreign Ministry that the international inspectors to visit district near the presidential palace in Baghdad as a "provocative step."

ملحق – 3

أبرز المؤسسات العاملة في مجال الترجمة الآلية من اللغة العربية وإليها إلى جانب "الحائك اللغوي".

1. مؤسسة صخر: هي المؤسسة الرائدة في مجال الدراسات والتجارب المتعلقة بإعداد اللغة العربية للترجمة الآلية. وتمتاز هذه المؤسسة بالاعتماد على جهود خبراء عرب ذوي معرفة باللغة العربية. ولها نشاط تجاري واسع فعال على الإنترنت. وتمتاز بتخصصها باللغة العربية. وكانت أنتجت برنامج "عجيب" بين الانكليزية والعربية، إلا أن نجاحه ظل محدوداً. وفي الفترة الأخيرة اتجهت اتجاهاً تجارياً، وركزت على الأغراض العملية مثل جهاز الترجمة للطبيب والاجهزة السمعية والبصرية والألعاب وغيرها.

2. سيستران Systran : لهذه المؤسسة نشاط واسع على الإنترنت، وعدة مواقع. كما أنها في حالة تطور دائم، ويعتبر برنامجها للترجمة الآلية V5 تطوراً مهماً في هذا المجال وتتجاوز جهودها اللغة العربية إلى مختلف لغات العالم. وفي شهر ايلول من عام 2006 قدمت معرضاً في القاهرة تركز حول النواحي التجارية.

3. سيموس CIMOS : تركز هذه المؤسسة على الترجمة بين العربية والفرنسية والانكليزية. وتعود جهودها إلى ربع القرن الماضي إذ كانت تستعمل برنامجها للترجمة الآلية في إدارتها المحلية وفي مبادلاتها التجارية، ومركزها الرئيسي في فرنسا ولها فرع في الولايات المتحدة. وقبل بضع سنوات انتقلت إلى المنافسة التجارية مع الشركات الأخرى العاملة في مجال اللغة العربية. وفازت بجائزة في مؤتمر التكنولوجيا اللغوية LANGITECH في تشرين الثاني/ نوفمبر 2003 .

والآن تعنى معظم المواقع العاملة في الترجمة الآلية على الإنترنت بإدراج اللغة العربية في قائمة لغاتها العاملة. وبرز مجدداً نشاط مؤسسة غوغل Googlo ، ونالت تقديرات عالية في اختبارات التقييم في الولايات المتحدة.

NIST 2005 Machine Translation Evaluation Official Results

Date of Release: Mon, Aug 1, 2005
Version 3

Track Data *Large* ,Arabic-to-English Task

Table 1	
Site	**BLEU-4 Score**
GOOGLE	0.5131
ISI	0.4657
IBM	0.4646
UMD	0.4497
JHU-CU	0.4348
EDINBURGH	0.3970
SYSTRAN	0.1079
MITRE	0.0772
FSC	0.0037

Data Track *Unlimited* ,Arabic-to-English Task

Table 2	
Site	**BLEU-4 Score**
GOOGLE	0.5137
SAKHR	0.3403
ARL	0.2257

Undefined Data Track ,Arabic-to-English Task

Table 3	
*LINEARB	*0.4300

automatic MT system. Linear B was not a submission of a fully * used non- Rather, it was a human-aided statistical MT system that Arabic speakers to correct the English fluency by selecting English phrases from the system's lattices. Search engines optional .the spelling of proper names were used to look up

By having humans in the loop and using a search engine at the time of evaluation, the Linear B system had access to data past the training cut-off date (December 1, 2004). Therefore, this system did not belong to either the" Large Data Track" or the "Unlimited Data Track" and is not directly comparable to the other systems in either track