اللغة العربية و الترجمة الآلية

سهيلة بربارة

معهد الترجمة – جامعة الجزائر2

Abstract

" The Arabic language and Machine Translation"

The world- wide technological breakthrough fostered by the development of computer science and digital media technologies, has completely reshaped the concept of translation that was a human task. It has become associated with the concepts and methods of computer-processing which has gained importance in its practice. The idea of designing a translating machine goes far back in history and has tapped into the dreams of the old researchers and hence, inspiring contemporary scientists since the twentieth century. The emergence of the first generation of computers in the wake of World War II, has paved the way for the application of a word- based Machine Translation (MT).

Then, research fields in this area have diversified to include the representation of linguistic theories through artificial intelligence applications in order to access Natural Language Processing. This resulted in the emergence of an interdisciplinary field called Computational Linguistics that enabled the design of more sophisticated systems for machine translation using mathematical algorithms. Since Machine translation systems are designed in the English language they have been better applied on languages of the same language family and have become the benchmark for progress and modernization.

This research deals with the analysis of the multifarious factors that made Arabic language lag behind in the field of Machine Translation as well as the problems of its computerization and translation toward it.

ملخص

أحدثت الطفرة التكنولوجيا عبر العالم و التي تعزّزت بتطوّر علم الحاسوب و تقنيات الإعلام الرقمية، إعادة صياغة كاملة لمفهوم الترجمة التي كانت حكرا على العنصر البشري، و التي غدت تقترن بمفاهيم وأساليب جعلت الحاسوب يتبوّأ مكانة هامة في ممارستها. إنّ فكرة تصميم آلة تُتَرجم تضرب في جذور التاريخ و لقد داعبت أحلام الباحثين القدامى و استوحى منها العلماء المعاصرون منذ القرن العشرين. فسحت نشأة الجيل الأول من الحواسيب غداة الحرب العالمية الثانية ، المجال لتطبيق الترجمة الآلية القائمة على المفردة.

تشعّبت، بعد ذلك، ميادين البحث في هذا المجال لتشمل مبادرات تمثيل النظريات اللسانية بفضل تطبيقات الذكاء الاصطناعي بهدف الوصول إلى المعالجة الآلية للّغة الطبيعية. نتج عن ذلك نشأة علم اللسانيات الحاسوبية المتعدّد التخصصات الذي مكّن من تصميم منظومات أكثر تطورا للترجمة الآلية باعتماد خوارزميات رياضية محدّدة. لما كانت لغة تصميم تلك البرمجيات هي اللغة الانجليزية ، فقد عرفت رواجا كبيرا في الدول الغربية وأصبحت معيارا للتقدّم والعصرنة.

يتناول هذا البحث تحليل عوامل تخلّف اللغة العربية في مجال الترجمة الآلية وإشكاليات حوسبتها و الترجمة إليها.

الكلمات الدالة:

علم الحاسوب، الترجمة الآلية، الذكاء الاصطناعي، المعالجة الآلية للّغة الطبيعية، اللسانيات الحاسوبية، البرمجيات، حوسبة اللغة العربية.

مقدمة:

يقودنا التطرق إلى هذا الموضوع المستحدث بالنسبة لأمّتنا العربية، إلى التعريف بمفاهيم الذكاء الاصطناعي (AI) (Artificial Intelligence) والمعالجة الآلية للغات الطبيعية (NLP)(Natural Language Processing) و اللسانيات الحاسوبية (Computational Linguistics) التي سمحت بنشأة الترجمة الآلية.

المعالجة الآلية للغات :

معالجة اللغات الطبيعية هي مقاربة حاسوبية لتحليل النّصوص. لعلّ أقرب تعريف لها هو أنّها علم فرعي من علوم الذكاء الاصطناعي الذي بدوره مُتفرّع من علوم الحاسوب أو المعلوماتية، ويتداخل بشكل كبير مع علوم اللّسانيات التي تقدّم التوصيف اللغوي المطلوب للحاسوب. هذا العلم يُمكِّننا من صناعة برمجيّات تتمكّن من تحليل فهم اللغات الطبيعية ومحاكاتها.

يهدف فهم الآلة للغة الطبيعية إلى تحليل اللغات الطبيعية على المستوى الصرفي والنحوي والدلالي والمقامي والسياقي.

الذكاء الاصطناعي:

وضع عالم الحاسوب الأمريكي جون مكارثي (John Mc'Carthy 1990) تسمية مصطلح الذكاء الاصطناعي لأوّل مرة عام 1956 و عرّفه على أنّه "علم وهندسة صنع الآلات الذكية" ، في حين يتفق باحثون آخرون على أنه دراسة الملكات العقلية للإنسان باستخدام النماذج الحسابية.

لعل أولى المحاولات في مجال الذكاء الاصطناعي هي الاختبار الذي ابتدعه آلان تيورنغ Alan Turing في 1950 للتأكّد من ذكاء الآلة عن طريق وضع الآلة في حجرة مغلقة تخرج منها نهاية طرفية في ردهة، ووضع إنسان في حجرة مغلقة أخرى يتّصل هو الأخر بنهاية طرفية في نفس الردهة، ويوجد إنسان آخر(الحكم) يتولىّ الاتّصال بالآلة والإنسان الأوّل وإدارة حوار مع كلّ من الآلة وقدرتها على التفكير و مدى نجاحها في خداع الحكم في حالة تكون أجوبتها مشابهة لأجوبة الإنسان.

لقد أطلق تيورنغ على ذلك الاختبار تسمية "اختبار التقليد" Imitation test أي بمعنى، تقليد الآلة للإنسان في الإجابة عن أسئلة معيّنة. وقد شكل اختبار تيورينغ منطلقا جديدا لأبحاث الذكاء الاصطناعي و تطبيقاته التي شملت الترجمة الآلية بمقاربة جديدة بعد أن كانت التطبيقات في هذا المجال تقتصر على بناء قواميس ثنائية اللغة على أساس ترجمة مفردة بمفردة و لقد سار على أثار تيورينغ ابتداء من الخمسينات، الثنائي الأمريكي(أكمان،1973، ص 452) آلان نيوول[1] Allen Newell وهربرت سايمن2 Herbert Simon اللذان وضعا برنامجين للذكاء الاصطناعي في 1956 و1957 و حازا في 1975 على جائزة تيورنغ.

1. آلان نيوول: (1927ـ 1992) هو باحث أمريكي في علم الحاسوب و في علم النفس الإدراكي، وضع لغة معالجة المعلومات في 1956 ، و كان أول واضع لبرنامجين في الذكاء الاصطناعي، الأول عام 1956 و الثاني عام 1957 بالاشتراك مع هاربرت سايمن .Herbert Simon

2. هربرت سايمن: (1916ـ 2001) هو باحث أمريكي في علم الاقتصاد و الاجتماع ، اهتم بعلم النفس الإدراكي و ثم بالذكاء الاصطناعي و يعتبر رائدا بمعية آلان نيوول في مجال البحث في الذكاء الاصطناعي.

الترجمة الآلية:

الترجمة الآلية هي نقل لنص أصلي مكتوب أو منطوق في اللغة لمصدر إلى نص مكتوب أو منطوق في اللغة الهدف باستخدام الحاسوب أي أنها تحويل من النص المصدر إلى النص الهدف باستخدام برمجيات حاسوبية.

تتحقق الترجمة الآلية بتوفر معرفة وخبرة في حقلين مختلفين هما "علم الحاسوب" و"علم اللغة أو اللسانيات" ويعد الرابط فيما بين هذين الحقلين،علم اللّسانيات الحاسوبية أو اللغويات الحاسوبية المستمدة أساسا من نظرية النحو التحويلي – التوليدي لنعوم تشومسكي و نظرية سيباويه النحوية (خشايمية، 2010) و التي تمّ تمثيلها حاسوبيا بواسطة خوارزميات رياضية محدّدة لفهم وتحليل اللغات الطبيعية.

تعود الإرهاصات الأولى في مجال البحث في الترجمة الآلية، ولا سيما ابتكار آلة تُترجم، كما نستخلصه من دراستنا لهذا الموضوع ومن منشورات الباحث الأمريكي المختص في مجال الترجمة الآلية جون هاتشنس (John Hutchins,2001,2)، إلى حقبة الثلاثينيات حيث بادر المهندس الروسي سميرنوف بتر بتروفيك ترويسنكيSmirnov Petr Petrovic Trojianski في 1931 بالتفكير في مشروع آلة تترجم و حاز على براءة الاختراع من أكاديمية موسكو للعلوم عام 1933 ، لكن المشروع لم يتحقق لعدم إمكانية تطبيقه على الآلات الحاسبة آنذاك. في الفترة ذاتها، حقق المهندس الروسي الأصل جورج أرترسوني Georges Artrsouni الذي هاجر للعيش في فرنسا حيث ابتكر ما أسماه بالدماغ الميكانيكي في 1932، أي آلة تترجم باسترجاع بيانات أكثرها حسابية استعملت لطباعة الكشوف المصرفية والجداول الزمنية للسكك الحديدية وأعقب ابتكاره بآلة ثانية مهّدت لبعث المساعي في مجال الترجمة الآلية.

لم تبرز الترجمة الآلية كموضوع بحث حول علم قائم بذاته إلا بعد حلول عهد الحواسيب وتصميم أول حاسوب الكتروني عام 1946 في الولايات المتحدة الأمريكية أُطلق عليه اسم "إنياك" ENIAC[1] ، فكان أول عمل بحث في مجال الترجمة الآلية هو المذكرة القيّمة الموسومة "الترجمة" التي حرّرها وارن ويفر(Warren Weaver) (1949، المختص في الرياضيات والذي كان يشغل منصب نائب رئيس مؤسسة الشهير روكفيلر والتي نُشرت عام 1949 لتشكّل انطلاقة الأبحاث في تطبيقات الترجمة الآلية بالجامعات الأمريكية.

تمّ عرض أول نظام للترجمة الآلية على الجمهور في عام 1954 في نيويورك على جهاز تابع لشركة آي بي أم IBM بمبادرة معهد مساتشوستس Massachussets Institute بجامعة جورج تاون الأمريكية). كانت تلك المنظومة تسمح بترجمة 49 عبارات روسية تم اختيارها بعناية إلى الإنجليزية باستخدام قاموس من 250 كلمة و ست قواعد نحوية وكانت المبادرة بدوافع استخبارية وعسكرية في ظل الحرب الباردة بين الكتلتين الشرقية والغربية آذاك. في 1955، انعقد أول مؤتمر للترجمة الآلية في الاتحاد السوفيتي سابقا وأعقبته مؤتمرات ومشاريع تصميم حواسيب لا سيما في بريطانيا، وكندا، وألمانيا وفرنسا.

بعد أن عرفت الأبحاث انطلاقة واعدة، جاء تقرير اللجنة الاستشارية للمعالجة الآلية للغة في الولايات المتحدة الأمريكية "ألباك")، (هاتشينس، 1996، ص ص9 -12) ALPAC في 1964 ليقضي بعدم جودة الترجمة الآلية بالمقارنة بالنتائج المحصل عليها و بإيقاف الدعم للمشاريع المتعلقة بها والتركيز فقط على الأدوات المساعدة على الترجمة

1. Electronic Numerical Integrator Analyser and Computer.

الآلية. لكن، مع تطور أنظمة الحواسيب وقدرتها ، انبعثت الأبحاث من جديد في الدول المذكورة آنفا وتم تصميم سستران Systran من قبل بيتر توما Peter Toma، وهو علم حاسوب فرنسي ومجري الأصل، كأول منظومة للترجمة الآلية في فرنسا عام 1968 وامتد الاهتمام بمجال الترجمة الآلية إلى أوروبا ولا سيما اليابان منذ السبعينات إلى يومنا هذا.

تجدر الإشارة إلى أنّ وتيرة التقدم المُحرز عليه في مجال الترجمة الآلية، كانت دوما مرتبطة بوتيرة تطور قدرة الحواسيب على استيعاب البيانات وعلى سرعة تنفيذها للعمليات المعقّدة، كما كانت مرتبطة أيضا بتطور النظريات اللسانية التي شكّلت القاعدة الأساسية لتصميم نظمها وتطور أساليبها.

أنواع الترجمة الآلية (بربارة، 2006، ص ص 38- 42)

هناك نوعان رئيسيان لأساليب الترجمة الآلية

- **الترجمة الآلية الخالصة أو الصرفة: Machine Translation**

تتمثل في تنفيذ منظومة الترجمة المدمجة في الحاسوب من النص المدخل Intput إلى النص المخرج Output بدون تدخل العنصر البشري.

- **الترجمة بمساعدة الآلة: (Machine Aided Translation M A T)**

هي التّرجمة التي تتمّ بمساعدة الآلة إلى الإنسان المترجم أو العكس و تقسّم بدورها إلى قسمين، كما يأتي:

أ– مساعدة الإنسان في الترجمة الآلية (Human Assisted Machine Translation- HAMT): هي المنظومات التي يتولى فيها الحاسوب مسؤولية ترجمة النص المدخل مع السماح للإنسان المترجم بالتدخّل من

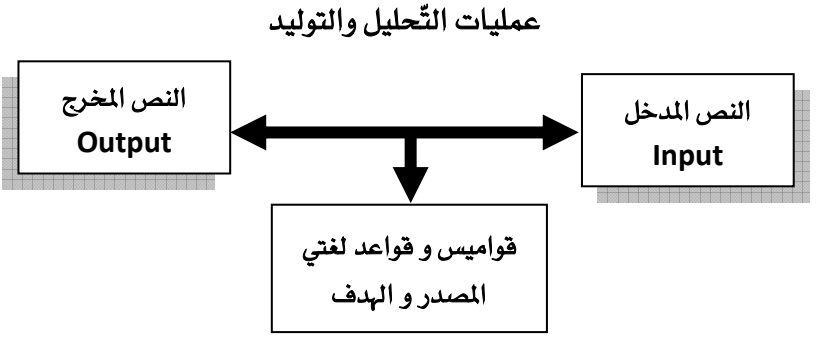
خلال شاشة الحاسوب و في أيّ مرحلة من مراحل الترجمة لمساعدة الحاسوب مثلا من أجل فكّ الغموض أو اللّبس لبعض الكلمات أو العبارات؛ وبذلك، فإنّ هذا الأسلوب يتيح للإنسان التفاعل مع الحاسوب للوصول إلى اللغة الهدف.

ب– مساعدة الآلة للإنسان المترجم (Machine- Assisted Human Translation) : يركّز هذا الأسلوب على قيام الإنسان المترجم بإجراء عمليات الترجمة وصولا إلى اللغة الهدف وذلك من خلال قيام منظومة الحاسوب بتوفير القواميس وعرض أمثلة لاستخدام بعض الكلمات والعبارات.

مناهج الترجمة الآلية:

المنهج المباشر : هذا الأسلوب هو الأول الذي اعتُمد في الجيل الأول من الحواسيب و كان يُستعمل بالنسبة لأزواج مُحدّدة من اللغات وهو يتضمن أساسا مقاربة الترجمة الآلية القائمة على المفردة (Word-based MT Approach) مع إدخال بعض التعديلات النحوية و يمكن توضيحها حسب الشكل التالي:

مراحل الترجمة الآلية باعتماد الإستراتيجية المباشرة

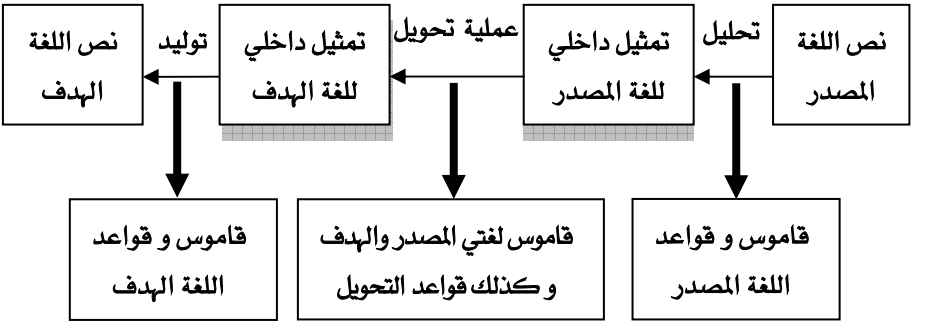


لأسلوب غير المباشر: تمتاز المنظومات التي تعتمد هذا الأسلوب على تحليل لغة المصدر وتوليد لغة الهدف بعمليات منفصلة، حيث يتمّ تحليل لغة المصدر بغضّ النّظر عما هو مطلوب توليده في لغة الهدف؛ ولهذه الإستراتيجية أسلوبان هما:

أسلوب التحويل للترجمة الألية (Transfer based MT): في هذا الأسلوب، يكون التمثيل الداخلي للوحدة التركيبية من النص (عبارة أو جملة) معتمدا على اللغة المصدر واللغة الهدف، وهذه تحتاج إلى مرحلة تسمىّ مرحلة التحويل التي تتمّ من خلالها مطابقة التمثيل الدلالي للغتي المصدر والهدف.

إنّ هذا الأسلوب يأخذ عبارات أو جمل اللغة المصدر حيث تُجرى عليها عمليات التحليل التركيبي و الخروج بتمثيل يسمى تمثيل اللغة المصدر. يعقب ذلك مرحلة التحويل والتي تجرى على مستوى القواميس والقواعد للحصول على تمثيل آخر مناسب يسمىّ تمثيل اللغة الهدف. أمّا المرحلة الأخيرة، فهي عملية توليد نص اللغة الهدف من التمثيل الداخلي للغة المصدر، والشكل الآتي يوضح ذلك :

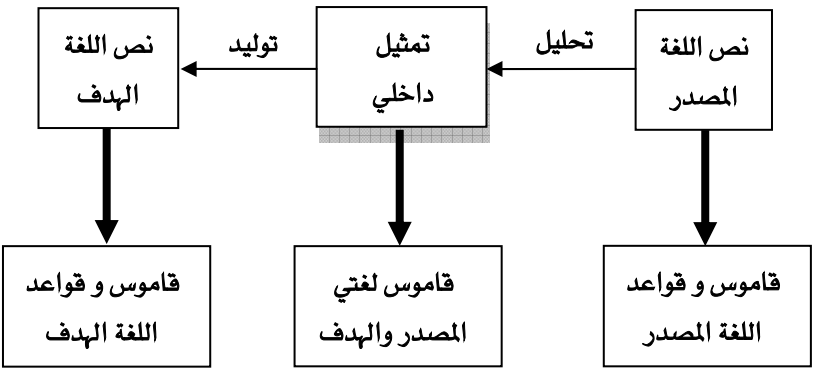
مراحل الترجمة الآلية باعتماد أسلوب التحويل في الإستراتيجية غير المباشرة



أسلوب الترجمة الآلية القائم على اللغة الوسيطة (Interlingua MT):

في هذا الأسلوب ، يحل مرحلة التحويل لغة وسيطة إذ يتم تمثيل داخلي وسيط بغضّ النظر عن لغتي المصدر و الهدف، حيث أنّ التمثيل الداخلي المستخرج من اللغة المصدر يمكن استخدامه لتوليد أي نص لغة أخرى (اللغة الهدف) وتتمّ عملية الترجمة بهذا الأسلوب من خلال مرحـلتين مختلفتين: الأولى يتمّ خلالها تحليل نص لغة المصدر والحصول على تمثيل وسيط له (Interlingual représentation)، أمّا في المرحلة الثانية، فيتم استخدام هذا التمثيل الوسيط لتوليد لغة المصدر، والشكل الآتي يوضح ذلك :

مراحل الترجمة الآلية باعتماد أسلوب اللغة الوسيطة
في الإستراتيجية غير المباشرة



اللغة الوسيطة هي لغة حاسوبية رياضية و قد تستعمل بعض منظومات الترجمة الآلية لغتين وسيطتين حيث تقوم الأولى بتحليل وتمثيل اللغة المصدر وتمر إلى اللغة الثانية التي تتولى تحليل مستويات وخصائص اللغة الهدف.

مقاربات الترجمة الآلية :

تتباين مقاربات الترجمة الآلية وفقا للأساليب المنتهجة في تصميم منظوماتها وهي تشمل ما يلي:

– **مقاربة الترجمة الآلية القائمة على قواعد اللغة** (علاونه، 2011) Rule- Based Machine Translation (RBMT) : هي مقاربة تقليدية إذ تمثل النظم الأولى للجيل الثاني من الحواسيب وهي تعتمد على القواعد اللغوية المرتبطة باللغات المصدر واللغات الهدف والمستخلصة أساسا من القواميس الأحادية اللغو أو الثنائية اللغة أو المتعددة اللغات وكذلك من المعاجم المتضمنة القواعد الأساسية، وهي تعتمد على الأسلوب المباشر في تصميم القواميس الالكترونية على سبيل المثال وعلى الأسلوب غير المباشر وأسلوب اللغة الوسيطة.

تشمل هذه المقاربة نظام الترجمة الآلية القائمة على التسلسل الهرمي للجمل (Hierarchical phrase-based translation) وهي تجمع بين نقاط قوة الترجمة القائمة على القواعد النحوية للغة و الترجمة القائمة على الجملة.

– **مقاربة الترجمة الآلية التجريبية** (ليماريز، 2008، ص ص 7– 9) (Empirical – Based Machine Translation (SMT): تعتمد هذه المقاربة على الذخيرة اللغوية أي على ذاكرة الترجمات خلافا للمقاربة السابقة ، حيث أنه يتم تعبئة الحاسوب بترجمات سابقة في عدة أزواج من اللغات يسترجعها المُترجم عند الحاجة وتصميم المنظومات بهذه المقاربة يستدعي تطبيق خوارزميات رياضية محدّدة وفقا للمنظومة المُعتمدة، مثل نظرية الاحتمالات ونظرية بايز الشهيرة (Bayes Theorem) ونماذج ماركوف الخفية (Hidden Markov Models) وما إلى ذلك، وتسعى جلّها إلى الحصول على النموذج الأنسب للترجمة.

تشمل هذه المقاربة نظما متباينة للمقاربة الإحصائية، نذكرها فيما يلي:

- **الترجمة الآلية القائمة على الأمثلة** (Example- based MT translation)

مقاربة الترجمة الآلية القائمة على الأمثلة (EBMT) هي أسلوب الترجمة الآلية يعتمد على متون تشمل نصوص موازية يتم إحضارها في عملية الترجمة. هي أساسا ترجمة عن طريق القياس أو التناظر، ويمكن أن ينظر إليها على أنها تطبيق لمقاربة الاستدلال المبني على دراسة الحالات لعملية تعلّم الآلة.

- **الترجمة الآلية الإحصائية** (كوستا جوس، 2012، ص 248) (Statistical Machine Translation) (SMT):

هي مقاربة من مقاربات الترجمة الآلية حيث يتم توليد الترجمات على أساس النماذج الإحصائية التي تعتمد على تحليل مجاميع من النصوص الموازية في زوج أو أكثر من اللغات وهي تتعدى معالجة الجمل والأمثلة إلى فقرات بأكملها .

أدخل وارن ويفر (Warren Weaver) الأفكار الأولى عن الترجمة الآلية الإحصائية ولو أنّ تطبيقها اقتصر على إحصاء المفردات ومقابلها في اللغة الهدف. أعاد الباحثون تطبيقها في أواخر الثمانينات و أوائل التسعينات من آي بي أم (Corporation International Business Machines) (IBM) ـ المؤسّسة الأمريكية الدولية للحاسبات الآلية في مركز البحث توماس واطسون[1]، وساهم في إثارة الاهتمام من جديد بالترجمة الآلية في السنوات الأخيرة. في الوقت الحاضر ـ يعتبر هذا الأسلوب من الترجمة الآلية من الأساليب الأكثر دراسة على نطاق واسع.

1. رئيس مؤسسة آي بي أم.

يعتبر فرانز جوزيف أوك[1] (Franz Joseph Och) من الأوائل الذين انتقدوا الخوارزميات المصمّمة لنظم الترجمة الآلية القائمة على قواعد اللّغة وأول من وضع المقاربة الإحصائية بالاستناد إلى الخوارزميات الرياضية كنظرية بايز للاحتمالات.

لعلّ منظومة غوغل للترجمة هي أحسن مثال عن المقاربة الإحصائية منذ أن أعلنت الشركة في 2006 عن ملكيتها لخوارزمية رياضية لتطبيق هذه المقاربة بعد أن كانت منظومة غوغل قائمة على نظام سستران الذي كان يعتمد المقاربة القائمة على قواعد اللغة.

– **مقاربة الترجمة الآلية الهجينة** (كوستا جوس و آخرون، 2016، ص ص 67- 75) (Hybrid Machine Translation Approach): **هي مقاربة تسعى للجمع بين مواطن قوة المقاربة اللغوية والمقاربة الإحصائية وتعتمدها حاليا بعض المنظومات مثل سستران وهي قيد التحسين.**

أمام النجاحات التي أحرزتها الترجمة الآلية ولا سيما بعد إنشاء الشبكة العالمية ، أصبح اللجوء إلى هذه الوسيلة أمرا لا مناص منه في عصر المعلومات الرقمية و انفجار الكم الهائل من المعارف وضرورة احتوائها في اللغة الأصلية.

في ضوء ما سبق، وبالنظر إلى تخلف العالم العربي في شتى مناحي الحياة، بات من الضروري الوقوف على محل اللغة العربية من الترجمة الآلية وإشكاليات الترجمة إليها على وجه التحديد و الطرح الذي يتبادر إلى الأذهان هو : كيف يمكن تفسير عزوف العرب عن حوسبة اللغة العربية؟ أيجدر التّسليم بعجزها، حسب اعتقاد هؤلاء، عن أن تكون

1. هو عالم ألماني مختص في البحث في علم الحاسوب، لقد ترأّس فريق البحث في مجال الترجمة الآلية في مؤسّسة جوجل .

موضع توليف و تطويع آلي ناجع؟ ما هي الحوائل التي تقف حاجزا أمام تصميم برمجيات عربية على غرار البرمجيات الغربية؟

إن الإجابة على هذه الطروحات تستدعى تحديد الإشكاليات التى حالت دون إعارة اللغة العربية العناية الكافية للحوسبة الآلية لأغراض الترجمة على وجه الخصوص وبالنظر إلى الحاجة الملحة التى أصبح الباحثون يلمسونها لإعادة بعث حركة الترجمة في البلاد العربية وضرورة اقترانها بالأدوات التكنولوجية المتاحة التى يشهدها عصر العولمة والشبكة الرقمية، فلقد أضحى من نافلة القول أن حوسبة اللغة مهمة حضارية لازمة لتطور اللغة ولبقائها[1] على حد تعبير الأستاذ مأمون الحطاب.

إشكاليات الترجمة إلى اللغة العربية (فرغلى و آخرون، 2012، ص ص 343- 347) :

باستثناء شركة صخر التي أنشئت عام 1982 كشركة كويتية تابعة لشركة "العالمية" للإلكترونات ثم بيعت الشركة و تمّ نقل مقرها إلى القاهرة ، والتي قامت بتطوير جيلٍ جديدٍ من تقنيات *المعالجة الطبيعية للغة العربية* و بتصحيح الاعتقاد الخاطئ بأنه يمكن تطويع الحلول المطوّرة في الغرب لتناسب احتياجات المستخدمين العرب ، برزت شركة أي تي آي (ATA) التي تأسّست في 1992 كالشركة المنتجة الرئيسية لبرمجياتٍ تستهدف اللغة العربية واللغة الإنجليزية في مجال الترجمة الآلية وهي تعتبر نفسها كالشركة الرائدة في العالم في الترجمة الآلية من اللغة الإنجليزية إلى العربية ولديها عدد كبير من المنتجات من مواقع وبرمجيات ترجمة آلية مثل مجموعة برمجيات "الوافي" وأدوات مساعدة على الترجمة.

1. مدير عام دار حوسبة النص العربي عمان – الأردن.

على الرغم من الجهود المؤسساتية والفردية على حد سواء، تمثلت المشاكل التي طالت حوسبة اللغة العربية فيما يلي:

-- **الخلفية التاريخية**: قبل عرض المشاكل العملية للترجمة إلى اللغة العربية، يجب الإشارة إلى أنه في الفترة التي شهدت انطلاقة الأبحاث في مجال الترجمة الآلية، كانت معظم الدول العربية تعاني ويلات الاستعمار والتراجع الاقتصادي والاجتماعي ودمار مؤسّساتها على يد الاستعمار الجائر الذي عمد إلى طمس هويتها والتشكيك بقوة لغتها في التعبير عن مظاهر ثقافتها وأصولها، و قد نجح إلى حد ما في ذلك. فقد نجح في زرع الذهن العربي فكرة عجز هذه اللغة عن أن تكون معيارا للتقدم والعصرنة. على الرغم من جهود بعض الدول العربية الرامية إلى تعريب المعارف، باءت تلك الحركة بالفشل بالحكم على اللغة العربية، عن جهل، بالضعف وتمّ حصر استخدامها "المقدّس" في المقام الديني وفي المراسيم الرسمية والخطابات و مقررات التدريس.

-- **سبب ذاتي**: على غرار المترجمين الغربيين، ساد تخوّف شديد لدى المترجمين العرب وعلى مرّ سنين عديدة، من أن تحلّ يوما الآلة محلّ المترجم، فتضمحلّ المهنة وتندثر. بيد إن الواقع بدّد هذا الاعتقاد على أساس نتائج تقييم نِتاج الآلة التي تَبيّن أنه بعيد عن الدقة في الترجمة التي يُحقّقها المترجمين البشر، في حين أصبح يُنظر للآلة على أنّها الحليف الضروري للمترجم كوسيلة مساعدة على الترجمة لما تتّسم به من مزايا في توفير الوقت وسرعة الإنجاز وثراء ذخيرتها من بيانات بما لا يسعه العقل البشري.

المشاكل العملية :

-- **المشاكل العامة:** هي مشاكل تشترك فيها اللغة العربية مع غيرها من اللغات (بربارة، 2013، ص ص 7- 14) و هي :

- المشاكل المادية المتعلقة بالتكلفة الباهظة لإنجاز وسائط تخزين ضخمة لاستيعاب الكم الهائل من الذخيرة اللغوية وفي أقصى سرعة.
- المشاكل البرمجية المتمثّلة في إيجاد الخوارزميات المناسبة لتصميم منظومات ترجمة أكثر جودة بحيث تحتوي على جميع جوانب تمثيل اللغة وتحليلها، انطلاقا من تقنيات تحليل وتمثيل قواعد اللغة وتقنيات الإحصاء القائمة على خوارزميات رياضية مبنية على نظرية الاحتمالات كما أسلفنا ذكره في جانب من هذا البحث.
- لغة تصميم برمجيات الترجمة هي اللغة الانجليزية مما قوّض من فرص استعارة تقنياتها لتصميم برمجيات في اللغات التي لا تنتمي إلى عائلة واحدة من اللغات، بما في ذلك اللغة العربية التي تختلف خصائصها عن غيرها من اللغات.
- ج ـ مشكلة اللبس والغموض التي تنطوي عليه اللغات على مستويات عديدة.

-- **المشاكل الخاصة:** تجدر إلى أن المشاكل الخاصة التي طالما شكّلت عائقا أمام حوسبة اللغة العربية و التي تمت تسويتها بنجاح بارز بفضل جهود تعريب أنظمة تشغيل الحواسيب و إدراج الحروف العربية في لوحة المفاتيح، كانت تتمثّل فيما يلي:

- مشكلة نظام كتابة اللغة العربية من اليمين إلى اليسار والتي لم تعد قائمة.

- مشكلة إدراج الحروف العربية و علامات التشكيل في لوحة المفاتيح، فقد أضحت المسألة أمرا يسيرا، بحيث أُدخلت الحروف العربية الثماني والعشرين في جميع أشكالها و حسب موضع كتابتها في الكلمة والهمزة بأشكالها والتاء المربوطة والألف المقصورة وأدوات التشكيل الثمانية والأرقام العربية و لألف الحنجرية وهمزة الوصل والهمزة العربية وتمّ إدخال اللام ألف جملة واحدة بل وبأشكال متعددة :لا لأ لإ لآ إلى جانب علامات التشكيل و الشدّة.

بالتالي، أصبحت اللغة العربية لا تعاني في حوسبتها من مشاكل مرتبطة بخصائص انتمائها وذلك بفضل تضافر بعض الشركات الغربية والسواعد العربية المُستخدمة فيها أو الشركات العربية النادرة المذكورة آنفا، بل يمكن حصر أبرز العراقيل التي تواجه نظامها اللغوي كأي نظام آخر على النحو الآتي بيانه:

مشكلة فك اللبس المعجمي: تتعدّد مواطن اللبس المعجمي بحسب طبيعته، ومن أهم مظاهره نذكر ما يلي:

- **اللبس الناشئ المرتبط بفئة الكلمات** Category ambiguity: في هذه الحالة يصعب للحاسوب التعرف على الفئة التي تنتمي إليها بعض المفردات العربية كما في الأمثلة التالية: **فرح، كتب، وعلم**، التي يمكن أتنتمي إلى فئة الأفعال أو الأسماء.

  الحل: يكمن في الاستعانة بالتشكيل الآلي Automatic diacritizer و المحلّل النحوي للفصل في وظيفة المفردة.

- **اللبس الناشئ عن الجناس اللفظي التام** Homograph ambiguity: الألفاظ المتجانسة هي أزواج أو مجموعات من الكلمات التي تُكتب

بالصفة نفسها، ولكن لها معان مختلفة، كما في الأمثلة العربية التالية: **المغرب، العشاء، العصر** التي قد تعني مواقيت الصلاة أو بالترتيب اسم البلد، ووقت العشاء، والمفردة الثالثة قد تعني وقت الصلاة، أو فترة زمنية.

- **الحلّ** : فك اللبس في هذه الحالة يمكن أن يتحقق بالاستعانة بالقواعد السياقية أو بالمقاربة الإحصائية القائمة على الأمثلة.

✓ **اللبس الناشئ عن مرجعية الضمائر** Pronoun reference ambiguity

الضمير هو كلمة تحل محل أو تأتي بديلا إما لاسم أو لعبارة اسمية، قد تُحدث ضمائر الإشارة الغموض عندما لا يكون واضحا الاسم الذي يشير إليه الضمير كما في ترجمة الجملة الانجليزية التالية إلى اللغة العربية:

المثال الانجليزي:

The **school fence** is old. Indeed, **It** was built in the last century

الترجمة إلى اللغة العربية:

- الاحتمال الأول: **سور المدرسة** قديم. بالفعل، فقد **شُيّد** في القرن الماضي.
- الاحتمال الثاني: **سور المدرسة** قديم. بالفعل، فقد **شُيّدت** في القرن الماضي.
- **الحل**: فك اللبس في هذه الحالة يمكن أن يتحقق بالاستعانة بالقواعد السياقية.

✓ **اللّبس الناشئ عن العدد ونوع الجنس**Gender and number ambiguity

اللغة الإنجليزية، مثل معظم لغات العالم، تُميّز فقط بين عدد المفرد والجمع. عندما نتحدث عن شخص واحد أو شيء واحد، يتم استخدام صيغة المفرد في حين تستخدم صيغة الجمع عندما نتحدث عن أكثر من شخص أو شيء واحد. في المقابل، نجد أن اللغة العربية لديها صيغ المفرد، والجمع، والمثنى في الأسماء، والأفعال، والصفات، وما إلى ذلك كما يتجلى في المثال التالي:

الترجمة إلى اللغة العربية:

المثال الانجليزي:

Salim and **Lina** applauded a lot. **They** were very delighted

ترجمة منظومة غوغل: سالم ولينا **صفق** كثيرا. **كانوا** سعداء جدا.

الصواب: سليم ونادية **صفقا** كثيرا. **كانا سعيدين** للعرض.

الحلّ: فك اللبس في هذه الحالة لا يمكن أن يتحقق إلا بالاستعانة بالمقاربة القائمة على قواعد اللغة.

✓ **اللّبس النحوي** Syntactic ambiguity

تتكون الجمل على المستوى النحوي، تتكون من سلسلة من الكلمات تربطها علاقات نحوية أي وظيفية أو هرمية. الطريقة التي يتم بها ترتيب الكلمات بشكل هرمي لها دور هام في تحديد المعنى. يحدث اللبس أو الغموض لّما يوجد أكثر من طريقة لتحليل البنية الأساسية للجملة وفقا لقواعد اللغة المستخدمة في النظام.

المثال الانجليزي: The wonderful paintings and sculptures

ملاحظة: يكمن اللبس في الصفة "wonderful " و هل تصف موصوفا واحدا أو الاثنين معا؟

الحلّ: فك اللّبس يحتاج إلى معرفة سياقية.

- ✓ **اللبس الناشئ عن ترتيب الكلمات في الجملة**

ترتيب الكلمات في الجملة الإنجليزية ثابت ويأتي في صيغة الجملة الاسمية: الفاعل + الفعل + المفعول به أو شبه الجملة الظرفية.

ترتيب الكلمات في الجملة العربية مرن يتأرجح بين **الجمل الاسمية**: فاعل+ فعل+ مفعول به

و **الجمل الفعلية**: فعل + فاعل+ مفعول به

مثال لجملة انجليزية لها ترجمة واحدة إلى اللغة العربية وأكثر من طريقة واحدة لترتيب الكلمات في الجملة:

Nihad has done her homework on the desk of the mistress in class

قامت نهاد بواجباتها على مكتب المعلمة في القسم.
في القسم قامت نهاد بواجباتها على مكتب المعلمة.
نهاد قامت بواجباتها في القسم على مكتب المعلمة.
على مكتب المعلمة قامت نهاد بواجباتها في القسم.

هناك أيضا اختلاف في موضع الصفة و الموصوف بين اللغة الانجليزية و اللغة العربية. ففي الانجليزية، الصفة تسبق الموصوف عكس اللغة العربية كما يتجلى في المثال التالي:

The beautiful sight and the fresh air

المنظر الجميل و الهواء المنعش.

✓ مشكلة اللبس الناجم عن التلازم اللّفظي

لعلّ أهم المعضلات التي تطرح نفسها في الترجمة الآلية بين اللغات و من ضمنها اللغة العربية، يجدر إثارة موضوع التلازم اللفظي ّ أو المتواردات أو المتصاحبات أو المترافقات، أو المقترنات كما يسمّيها البعض، تستخدم للإشارة إلى كلمة يقترن استخدامها بكلمة أو كلمات أخرى كما في الأمثلة العربية التالية: **صديق حميم، عدو لدود، حديقة غناء، شجرة باسقة، أموال طائلة، سماء صافية، أمطار غزيرة...**

هناك نوع آخر من المتلازمات اللفظية وهي **الأمثال والحكم التي تشكل** وحدات معنى لا يمكن تجزئة مفرداتها لأن علاقة التوارد فيما بينها هي علاقة وطيدة وشرطية، ولعل الأمثلة الآتية توضح ذلك: **وافق شنّ طبقة ، رجع بخفّي حنين، ربّ عذر أقبح من ذنب.**

إذا كانت ترجمة الأمثال والحكم والعبارات المسكوكة تشكّل معضلة بالنسبة للمترجم البشري، ناهيك عن معضلة معالجتها حاسوبيا.

الحل: اعتماد المقاربة الإحصائية يكون بدون شك حلا عمليا لتجاوز المشكلة.

✓ مشكلة اللبس الناجم عن ترجمة العبارات المختصرة و أسماء الأعلام و المؤسسات

غالبا ما يكون نتاج جودة الترجمة إلى اللغة العربية ضعيفا عند ترجمة العبارات المختصرة وأسماء الأعلام والمؤسسات، لأنها مشكلة يواجهها في الأصل المترجم البشري، وإن كان حلّ الترجمة البشرية كامن في وجوب إلمام المترجم بالكفاية المعرفية التي تؤهله لتفادي الأخطاء المحتمل الوقوع فيها في مثل هذه الحالات ، فإن الترجمة الآلية تبقى رهينة جودة البرنامج المصمّم لها. لنلاحظ أخطاء ترجمة

المختصرات والأسماء الانجليزية إلى اللغة العربية بواسطة منظومة غوغل والوافي الذهبي الصادر عن شركة ATA ، على سبيل المثال:

العبارة المختصرة باللغة الانجليزية: MAT
ترجمة غوغل: حصيرة
ترجمة الوافي الذهبي: الحصيرة
الصواب: الترجمة بمساعدة الحاسوب

الحل: إحصاء العبارة المختصرة حسب المجال وترجماتها إلى اللغة العربية وإدراجها ضمن الوحدة النمطية للترجمة الآلية.

مثال عن أسماء الأعلام والمؤسسات:
المثال الانجليزي: Mrs Bird Paddington
ترجمة غوغل : السيدة الطيور بادينغتون
ترجمة الوافي الذهبي: السيدة Bird Paddington

الصواب: السيدة بيرد بادينغتون التي هي شخصية لمدبّرة بيت في سلسلة تلفزيونية. يتعين في مثل هذه الحالة، كتابة الاسم بالأحرف العربية وإعادة كتابته باللغة الأصلية بين قوصين.

الحل: إحصاء أسماء الأعلام الأكثر شهرة أو على أساس المجال وإدراج القائمة ضمن الوحدة النمطية لمنظومة الترجمة الآلية.

المثال الانجليزي: WHO
ترجمة غوغل : من الذي
ترجمة الوافي الذهبي: من
الصواب: المنظمة العالمية للصحة.

الحل: إحصاء أسماء المؤسسات المعروفة أو على أساس المجال وإدراج القائمة ضمن الوحدة النمطية لمنظومة الترجمة الآلية.

خلاصة البحث:

في ضوء دراستنا لبدايات الترجمة الآلية وتطورها بصفة سريعة ومذهلة عند الغرب وللجهود التي بذلتها الأطراف العربية حتى وإن كانت متواضعة، إلا أنها ساهمت بشكل كبير بالتعاون مع الشركات العالمية والاستعارة من المناهج المعتمدة في هذا المجال، في تذليل بعض العقبات التي كانت تبدو مستعصية مما فتح آفاقا واعدة أمام الترجمة من اللغة العربية وإليها. فلقد مكن تصميم الأدوات المساعدة على الترجمة مثل المدقق الإملائي، والمحلل الصرفي والإعرابي، والتشكيل الآلي والتعرّف الضوئي على الحروف العربية وغيرها من الإنجازات الجليلة، من الارتقاء بالتحليل الحاسوبي السطحي إلى تطوير أدوات ناجعة في المستويات الأساسية، كالمحلل الصرفي، إذ أثبتت اللغة العربية قابليتها أن تُعالَج حسابيا كنظام منطقي يمكن أن تُستخلص منها القواعد والقوانين والشروط التي تكون الأسس والخصائص العامة لكل اللغات، على طريقة نعوم تشومسكي في نظره إلى اللغة على أنها نظام صوري.

في هذا الشأن، يكفي أن نقف على حقائق نُظم اللغة العربية في المستويات الصوتية والصرفية والنحوية و اختيار الخوارزميات الأنسب للتمثيل الحاسوبي لها ولتصميم برمجيات تمزج بين أدوات المعالجة اللغوية وبين المقاربة الإحصائية في وحدة نمطية واحدة لتحقيق المقاربة الهجينة للترجمة الآلية ونظام سستران أحس مثال عن المساعي المعاصرة الرامية للجمع بين مواطن قوة المقاربة الإحصائية بالاعتماد على ذاكرة الترجمة بصفة أساسية مع إدماج أدوات تحليلية للغة.

الاستجابة لهذا العرض تستدعي مشاركة جميع الأطراف الفاعلة من علماء حاسوب وعلماء لغة ومختصين في الرياضيات في مرحلة تصميم

منظومة الترجمة والاعتماد على المترجمين المحترفين في مرحلة تقييم جودة النتاج، لأن الجهات الغربية التي تقوم بحوسبة اللغة العربية في الوقت الحاضر، لا تعتمد على التحليل العميق لها وإنما تنتهج تحليلا سطحيا وتُدرجها في منظوماتها لأغراض اقتصادية وتجارية، لا تخدم لغة الضاد كثيرا.

على وجه العموم يجب أن يمنح أهل لغة الضاد للغتهم العناية الكافية من أجل استنباط القواعد التي تسيّرها بهدف تمثيلها بصفة يستوعبها الحاسوب.

يجدر لفت الانتباه إلى حقائق طالما أحجمت عن ترويجها الشركات العملاقة المنتجة للحواسيب وهي كما ذكره الأستاذ الدكتور محمّد بطّاز[1] مدير شبكات وأنظمة المعلومات والاتّصال الجامعيّة، ورئيس المجمع الجزائري للغة العربية، في محاضرة ألقاها بمجمع اللغة العربية الأردني في موسمه السادس في 2008 حول مسألة بناء الأجهزة الحديثة وفق خصائص اللّغة العربيّة إذ شرح قابلية اللغة العربية للحوسبة موضّحا أنه "من المعروف عن الحاسوب أنّه لا يفهم إلا لغة الأرقام وبالذّات الأرقام العربيّة، وهذا يعني بطريقة أدقّ أنّ الأجيال المتتالية للحاسوب ولغات البرمجة وشبكات الحاسوب والتّكنولوجيّات المختلفة المستخدمة لبنائها تعتمد أساسًا على نموذج جبريّ محض، وهذا يعني بكلّ بساطة أنّه لولا المدرسة العربيّة لما وجد الحاسوب الرّقميّ بالنّمط المعروف حاليًّا على الأقلّ" كما أكّد في باب آخر، في ندوة مداخلة له في ندوة علمية عام 2010 حول موضوع "اللغة العربية في مجال التشبيك

1. باحث و علامة جزائري، تقلّد عدم مناصب رفيعة من بينها عمادة كلية تكنولوجيا المعلومات بجامعة فيلاديفيا – عمان – الأردن، حاز على جائزة الملك فيصل عام 2010.

الالكتروني" بأنه "يجب أن ندرك أن الأمر ليس أمر لغة بقدر ما هو مسألة قدرة على الدخول في معركة الإنتاج وقدرة على تشريح المنتجات المتوفرة لتكييفها وجعلها تتوافق مع لغتنا وثقافتنا ورؤانا في عالمنا هذا. نقول أن الأمر ليس أمر لغة ونكرر هذا لأنه من المعروف عن الحاسوب أنه لا يتكلم إلا لغة الأرقام. وإذا أردنا أن ننسب صفة لهذه الأرقام فإننا سنقول الأرقام العربية وإذا أردنا أن نصف عمليات الحاسوب فإننا نقول عمليات جبرية. ولكن إذا نظرنا إلى البلاد العربية فإننا نجدها في الواقع تستهلك الكثير ولا تنتج إلا القليل من البرامج الحاسوبية وهذا بالرغم من البحوث المتطورة والمجهودات الجبارة التي يبذلها الباحثون العرب في هذا المجال بالذات".

بالتالي، باتت الحاجة ملحة لتنسيق الأعمال الفردية على اختلاف مجالات الاختصاص ذات الصلة بحوسبة اللغة العربية والتعريف بها على نطاق واسع للعودة باللغة العربية إلى المكانة التي كنت تتبوؤها في عصر الترجمة الذهبي الذي بنى عليه الغرب في وقت لاحق حضارتهم.

المصادر و المراجع:

أ- باللغة العربية

1. سهيلة بربارة (2013): *إشكاليات الترجمة الالكترونية*: محور المداخلة في الجلسة الثامنة من المؤتمر الدولي الأول حول "الترجمة وإشكاليات المثاقفة" في 26 و 27 فبراير 2013 بالدوحة، قطر بتنظيم منتدى العلاقات الدولية و العربية، و موضوع البحث المنشور في الكتاب الأول للمنتدى .
2. أ. فتحي خشايمية (2010) : *نظرية النحو التوليدي التحويلي عند تشومسكي.*
3. سعيد حسن بحيري (1989): *"عناصر النظرية النحوية في كتاب سيبويه: محاولة لإعادة التشكيل في ضوء الاتجاه المعجمي الوظيفي دراسات في علم اللغة التقابلي"* ، الطبعة الأولى ، مكتبة الأنجلو المصرية.
4. شريف خطاب (2011) : *"المعالجة الآلية للغة العربية، أساسيات الحاسوب"* ، بحث، قسم علوم الحاسب، كلية الحاسبات والمعلومات، جامعة القاهرة.
5. د. سعد بن هادي القحطاني : *" تحليل اللغة العربية بوساطة الحاسوب"* مركز اللغة الإنجليزية – معهد الإدارة/ الرياض.
6. د. نبيل على (1998) : *اللغة العربية و الحاسوب*. الكويت: مؤسسة تعريب الكويت.
7. سهيلة بربارة (2006): الترجمة بمساعدة الحاسوب من الانجليزية إلى العربية. مذكرة لنيل شهادة الماجستير، قسم الترجمة، جامعة الجزائر 2، ص ص 38 – 42).

ب ـ باللغة الأجنبية

1. McCarthy, J. (1990) : "*Generality in artificial intelligence*". In Lifschitz, V., ed., *Formalizing Common Sense*. Ablex.

2. Varol Akman (1995) : *Formalizing Common Sense: Papers by John McCarthy*, *Book Review Lifschitz, ed., (Ablex Publishing Corporation, Norwood, NJ, 1990.* Ed Elsevier, Artificial Intelligence, N° 77, 1995.

3. Donald M. Johnson (1973) : *Human Problem Solving* by Allen Newell, Herbert A. Simon, *Reviewed book.* The American Journal of Psychology, Published by: *University of Illinois Press,* Vol. 86, No. 2,p 452.

4. Jacqueline Léon (2001) : *Le traitement automatique des langues.* Revue électronique Histoire, Epistémologie, Langage. Vol. 23 (1), 2001: 7-31 (Ref : publication de l'article de John Hutchins intitulé *« Machine translation over fifty years »* University of East Anglia, Norwich, UK , P 2. (http://www.hutchinsweb.me.uk /HEL-2001.pdf)

5. Hutchins, W.J.(1987): *Machine Translation: Past, Present, Future*, Ellis Horword Limited, West Sussex <http:ourworld. compuserve.com/hompages/WJ Hutchins/PPF_3.pdf) (assessed 29 December 2005)

6. John Hutchins: *The history of machine translation in a nutshell:* Web: http://ourworld.compuserve.com/homepages/WJHutchins], Revision 2005.

7. Shannon, C. E. and Weaver, W. (1949): *The mathematical theory of communication.* Urbana: University of Illinois Press.

8. Richens, R.H. and Booth, A.D. (1955)**:** *'Some methods of mechanized translation.'* In: Locke, W.N. and Booth, A.D. (eds.) Machine translation: fourteen essays (Cambridge, Mass.: Technology Press of the Massachusetts Institute of Technology), pp. 24-46.

9. John Hutchins (1996): *ALPAC: the (in) famous report.* From the archives, [from: MT News International, no. 14, pp. 9-12].

10. Chris Callison-Burch (2007) **:** *Machine translation: Word-based models and the EM algorithm*.Slides borrowed from Philipp Koehn, John Hopkins University December 3, p .

11. Mouiad Fadiel Alaouneh (2011) : *Rule-Based and Example-Based Machine Translation from English to Arabic.* Fac. of Inf. Sci. & T.echnol., Nat. Univ. of Malaysia, Bangi, Malaysia. Published by EEE Conference Publications.

12. Jesus Angel Gimenez Limarez (2008) : *Empirical Machine Translation and its Evaluation.* Tesi Doctoral per a optar al grau deDoctor en Inform`atica. Programa de Doctorat en Intellig`encia Artificial, Departament de Llenguatges i Sistemes Inform`atics, Universitat Polit`ecnica de Catalunya, Barcelona, , PP 7-9.

13. Marta R. Costa-Juss (2012): *Study and comparison of rule-based and statistical Catalan- Spanish Machine Translation Systems, Computing and Informatics*, Vol. 31, P 248.

14. Marta **R. Costa-jussà** (and others) (2016): *Hybrid Approaches to Machine Translation*. eBook ISBN : 978-3-319-21311-8, pp 67- 75.

15. Marta **R. Costa-jussà** (2015): *Latest trends in hybrid machine translation and its applications*. Computer speech and language , volume 32, issue 1, , PP 3- 10. Paper on ScienceDirect website: http://www.sciencedirect.com/science/article/pii/S0885230814001077.

16. Abdelhadi Soudi, Ali Farghaly, Günter Neumann, Rabih Zbib (2012): *Challenges for Arabic Machine Translation* .John Benjamins Publishing, pp343 – 347.