مجلة فصل الخطاب Journal of Faslo el-khitab
ISSN:1071-2335/ E-ISSN:2602-5922/ Legal Deposit N°: 2012-1759
مجلد 11، عدد رقم: 04، ديسمبر 2022، صص: 417 - 432

تاريخ الاستلام (2022/10/16) تاريخ القبول (2022/11/22) تاريخ النشر (2022/12/30)

المجالات التطبيقية للمعالجة الآلية لمستويات اللغة العربية،تحديات و آفاق

Practical areas for the automatic processing of the levels of Arabic: Challenges and prospects

الهام بغداد

جامعة محمد الخامس (المغرب)، ilham_boughdad@um5.ac.ma

ملخص:

لما كان هدف اللسانيات الحاسوبية تمكين الآلة من الإنتاج والتوليد لا الاستدعاء والمحاكاة بالتعامل مع نصوص غير محددة تراكيبها سابقا ولا دلالتها وجب الوقوف عند البنية السطحية وملاحظة ما يتحقق فيها مع محاولة البحث عن دخائل هذه البنية في البنية العميقة، لأن مدخل التوليد هو فهم المنطق الذي يعمل به اللسان البشري، وفي ظل غياب هذا التصور سيظل الاستدعاء هو الغالب على مجال حوسبة اللسان العربي. من هذا المنطلق سنقدم في هذه الورقة لمحة عن تطبيقات المعالجة الآلية لمستويات اللغة العربية محاولين بيان أهم التحديات التي تواجهها في ظل غياب تحليل أولي لمستويات اللغة العربية، فما يعرفه مجال حوسبة اللغة العربية من عقبات مرده في غالبه إلى غياب تحليل منطقي للإشكالات التي يعاني منها اللسان العربي في كافة مستوياته.

الكلمات المفتاحية: الحوسبة، اللسانيات، اللغة العربية، الخوارزميات.

Summary:

Since the goal of computational linguistics is enabling the machine to produce and generate, not to recall and simulate, through processing texts whose structures and connotations are not previously predetermined, it is necessary to stop at the surface structure and observe what is realized in it while attempting to look for its components in the deep structure, because the field of generation leads to the

المؤلف المرسل: الهام بغداد، الإيميل: ilham_boughdad@um5.ac.ma

understanding of the logic lying behind the actualization of the human language. In the absence of this perception, recalling will prevail in the domain of Arabic computerization. From this outlook, we will present in this paper an overview about the applications of the automatic processing in the different levels of Arabic in order to depict the main challenges and obstacles this latter faces in the absence of a preliminary analysis of its levels. This absence is generally attributable to the lack of an analysis of the problems Arabic undergoes in all its levels.

Keywords: computing - linguistics - Arabic language - algorithms.

1. مقدمة:

ساهمت النقلة المعلوماتية التي عرفها العالم في ولوج اللغة العربية عالم الرقمنة بمعالجتها آليا، إذ حظيت بالعناية من قبل الدارسين العرب وغير العرب شأنها شأن باقي اللغات الطبيعية، وقد ساعد في ذلك ما تتميز به هذه اللغة من خصائص تساعد في حوسبة جميع مستوياتها خاصة المستوى الصرفي والصوتي الذي يُعرف بنظامه الرياضي المنطقي، وإذا كان هذا المستوى لم يطرح تحديات كبيرة أمام الدارسين على ما يبدو، فإن باقي المستويات شكلت تحديا كبيرا لهم خاصة فيما يرتبط بنظامها التركيبي النحوي والدلالي.

إن هدف اللسانيات الحاسوبية لا ينحصر في الاستدعاء والمحاكاة بل يتجاوز ذلك إلى تمكين الآلة من التوليد والإنتاج بالتعامل مع نصوص جديدة غير محددة تراكيبها سابقا ولا دلالتها، ولتحقيق ذلك يجب الوقوف عند البنية السطحية وملاحظة ما يتحقق فيها مع محاولة البحث عن دخائل هذه البنية في البنية العميقة قصد استنباط الخوارزميات المتحكمة في الفعل اللغوي، لأن مدخل التوليد هو فهم المنطق الذي يعمل به اللسان البشري، وفي ظل غياب هذا التصور سيظل منطق التقليد والمحاكاة غير الممنهجة هو الغالب على مجال حوسبة اللسان العربي.

2.اللغة العربية والحاسوب أية علاقة؟

لقد كانت الآلات الذكية وستظل حلم علماء الحاسوب ومهندسيه ، فمنذ ثلاثينيات القرن الماضي بشر عالم نظرية الأوتوماتيات الشهير "آلان تورينج ALAN TURING" بإمكانية إنتاج آلة تحاكي في قدرتها كثيرا من قدرات البشر الذهنية ، وخلال الأربعين سنة الماضية، ومنذ ظهور الحاسوب في نهاية الأربعينيات، وعبر أربعة أجيال منه ترمز إلى مراحل مميزة في تطوره التقني، «ظلت الحواسيب الآلية رغم كل ما حققته من إنجازات باهرة في جميع الميادين محصورة في قدرتها الهائلة على تخزين المعلومات وتنظيمها واسترجاعها، إلى جانب

سرعتها الفائقة في تنفيذ العمليات الحسابية»[1]، ومع انطلاق عصر المعلومات، تعاظم الدور الذي تلعبه القوى الرمزية، وعلى رأسها اللغة، في صياغة شكل المجتمع الإنساني الحديث، إذ أصبحت في حاجة إلى منظور ذو توجه جديد، يعيد النظر في جميع جوانب المنظومة اللغوية من أعمق الجذور إلى أدق الفروع لتسهل عملية صورنة الفعل الكلامي عبر ترميزه واختزاله في قواعد تسهل عملية تخزينه وكذلك معالجته آليا إن تطلب الأمر ذلك.

وتعد دراسة القدرة اللغوية للإنسان من أهم مجالات البحث في علم الذكاء الاصطناعي، فاللغة هي وسيلة للتفكير والتعبير عن الأفكار والمعلومات ونقل الخبرات إلى الآخرين، وتدخل محاكاة هذه القدرة اللغوية في البرمجيات اللغوية التي يضعها علماء الذكاء الاصطناعي كالأنظمة الخبيرة وتمثيل المعرفة والتعليم بمساعدة الكمبيوتر وبما أن الجزء الأكبر من المعرفة الإنسانية موجود في نصوص لغوية، فلا بد للكمبيوتر من قواعد تمكنه من التعامل مع هذه النصوص لاستخلاص هذه المعرفة بعيدا عن منطق الاستدعاء الذي يقوم على التخزين في ذاكرة الحاسوب لأن هذا لا يعتبر إنتاجا وإنما تقليدا للقدرة الكلامية الإنتاجية لدى الإنسان.

3.تطبيقات المعالجة الآلية لمستويات اللغة العربية.

إن دور اللغة لا يقتصر على مكون التواصل كما هو متداول في بعض الدراسات اللسانية، بل إن لها أدوارا أخرى، تتمثل في ترميز البيانات اللغوية ذات الطابع الذهني والمنطقي، «كما لها القدرة على تمثل الواقع والقدرة على التخطيط، والقدرة على الخلق والإبداع الذهني للعوالم الممكنة وغير الممكنة. فاللغة لها بنية معرفية وذهنية تجعلها قادرة على ترميز وتمثل العالم الفيزيائي الواقعي والاستعاري»[2]

وتنظر اللسانيات الحاسوبية إلى اللغة على اعتبار أنها فعل منطقي يحتاج إلى الصورنة للوصول إلى العلاقات المتحكمة في إنتاج الفعل الكلامي؛ لذلك سعت إلى معالجتها تبعا للمستويات الفاعلة في هذا الفعل، ولعل أول مستوى يتحكم في إنتاج اللغة المستوى الصوتي ويتأكد ذلك من خلال تعريف ابن جني للغة، إذ يقول:«حد اللغة أصوات يعبر بها كل قوم عن أغراضهم»[3]، لكن لا يجب أن نلخص اللغة على أنها أصوات فقط؛ فإذا أردنا أن بحث في أصل هذه الأصوات سنجد أنها تحققت بعد أن طبق الذهن مجموعة من القواعد الصرفية والتركيبية والدلالية بل والصوتية أيضا، فلو قصرنا اللغة في الصوت دون هذه المستويات الأخرى لأصبحت كل الأصوات المسموعة لغة.

1.3المستوى الصوتي:

تسعى اللسانيات الحاسوبية إلى إكساب الحاسوب مهارة تحويل المنطوق إلى مكتوب ومهارة الضبط والقراءة الصوتية للنصوص المدخلة سواء عن طريق لوحة المفاتيح أو عن طريق القارئ الآلي دون ارتكاب أخطاء إملائية بين الأصوات المتشابهة والمتقاربة في المخارج، و«هذه التقنية تستخدم في الإملاء الآلي والترجمة الآلية الشفوية، والحوار الشفوي مع الآلة، إلا أنه لم يتحقق منها إلا الجزء اليسير في اللغة العربية، بسبب عدم اكتمال بناء العتاد اللساني، وخصوصية العربية في اقتصارها على كتابة الصوامت دون الصوائت، والطابع الانصهاري لكلماتها في المستوى الصرفي، واعتماد حركة آخر الكلمة في المستوى النحوي».[4]، فكل هذه العوامل وغيرها تقف عائقا أمام حوسبة المستوى الصوتي للغة بمعزل عن باقي المستويات، خاصة وأن المستوى التركيبي/النحوي هو شكل المنطوق وفي ظل غياب علاقة التأثير والتأثر بين هذه المستويات تبقى معالجة اللغة صوتيا أمرا مستحيلا إلى حد ما.

ومن التطبيقات المهمة للمعالجة الآلية للغة تطبيق تحويل النص إلى كلام منطوق الذي يقوم بقراءة النصوص أو تحويل الكلام المكتوب إلى صوت مسموع وكلام منطوق مفهوم، و«يسمى نظام الحاسب "برمجيات أو أجهزة" المستخدم لهذا الغرض بمؤلف الكلام، ويقوم نظام تحويل النص إلى كلام بتحويل نص اللغة العادية إلى كلام، أما الأنظمة الأخرى فتعمل على تحويل الرموز اللغوية الصوتية إلى كلام. ويمكن إنشاء الكلام المؤلف من خلال ضم أجزاء متسلسلة من الحديث المسجل المخزن في قاعدة بيانات هي المكانز المنطوقة، وتختلف الأنظمة في حجم وحدات الحديث المخزنة، وفي مجالات استخدام معينة، يتيح تخزين كلمات كاملة أو جمل كاملة إنتاج كلام عالي الجودة. وهناك طريقة أخرى هي تضمين نموذج من جهاز النطق وغيرها من خصائص الصوت الإنساني لإنتاج صوت مؤلف بالكامل.»[5]، لكن تلك الخصائص لا يمكن أن تنتج لنا كلاما يماثل الكلام البشري في ظل غياب قواعد بناء الكلمة وهندستها، فما ينتجه المتكلم البشري لا يُنتج اعتباطا وإنما يُنتج عن طريق قواعد مضمرة منها ما هو مكتسب لذا المتكلم ومنها ما هو فطري تتحكم فيه قواعد اللسان الذي يتكلم به ومنها ما يرجع للأعضاء النطقية المسؤولة عن إنتاج الأصوات إذ يستحيل إنتاج كلمة تتقارب مخارج أصواتها "الحروف"، لأن البنية النطقية لدى الإنسان لا تسمح بذلك.

ومن التطبيقات المتوصل لها اليوم في مجال حوسبة اللغة تطبيق تحويل النص إلى كلام، وتجدر الإشارة إلى أن هذا التطبيق قد اتسع ليشمل القيادة الآلية للطائرات العسكرية والتحكم في أجهزتها و«خاصة في بريطانيا وفرنسا والولايات المتحدة وفي طائرات الهليكوبتر حيث مشكلة الضوضاء الخلفية بسبب صوت المروحية والهواء وفي إدارة المعارك حيث تتطلب مراكز القيادة الوصول السريع لقواعد بيانات المعلومات المتغيرة بسرعة كما تستخدم في تدريب مراقبي حركة المرور الجوي، بجانب استخدامها في مجال الاتصالات التليفونية وألعاب الحاسب والمحاكاة. ولم يتم تثبيت هذه التكنولوجيا في الأجهزة المحمولة لأنها تتطلب قوة معالجة هائلة، وهذه التكنولوجيا مفيدة للغاية لمن لا يستطيعون تحريك أيديهم، مما يتطلب وسيلة بديلة لإدخال المعلومات في الحاسب والتحكم في وظائفه»[6]، أما فيما يخص حوسبة الجانب الصوتي في اللغة العربية فإن شركة الهندسة لتطوير النظم الرقمية ساهمت في هذا المجال بتقديم تقنيات معالجة اللغة العربية المكتوبة والمتمثلة في المشكل الآلي للنص العربي"سيبويه"، والمحلل الصرفي العربي "الميزان"، والمعنون النحوي للمفردات العربية "السراج" والتحليل الدلالي المعجمي العربي"العين"، ومحرك البحث في النص العربي، ومن بين البرامج التي أنتجتها أيضا يمكن أن نذكر:

- تقنيات معالجة الكلام المنطوق والمتمثلة في مساعد التعليم التفاعلي لتجويد القرآن الكريم أو اللغة العربية المنطوقة "حفص" ونظام تخليق الكلام المنطوق من النص العربي المكتوب "البليغ" وضبط الكلام المنطوق.
- تلخيص النصوص العربية "سارة".
- القراءة الآلية للنصوص العربية.
- برنامج Natiq الاحترافي المتخصص في تحويل النص إلى كلام.
- برنامج Kateb للتعرف التلقائي على الكلام في الوقت الحقيقي.
- برنامج سطور للتعرف الضوئي على الحروف المكتوبة باللغة العربية[7].

2.3 المستوى المعجمي:

يكاد يتفق معظم الدارسين أنه ليس من المفيد الاكتفاء بالمعاجم القديمة الورقية الموضوعة على طريقة المناهج التقليدية «لأن تقنيات التخزين ومعالجة المعلومات التي توفرها الآلة تمكن من بناء معاجم آلية وفق ضوابط لسانية وحاسوبية صارمة، والوضع

الحالي للغة الضاد، يفرض أكثر من أي وقت مضى التفكير جديا في بناء معاجم إلكترونية للغة العربية على غرار اللغات الأجنبية، قصد تقليص هوة الفجوة الرقمية بين المعاجم العربية والمعاجم الأجنبية. إن المتنفس الطبيعي لكل النظريات والمناهج اللسانية أصبح هو ما تحققه المعلوميات من تقدم في صياغة البرامج القادرة على تقييس دماغ الإنسان، وبذلك أصبحت قادرة على صياغة قوانين صورية تقوم بدور مزدوج، من جهة وصف النظام اللغوي في سائر مستوياته باستعمال لغة عقلانية، ومن وجهة أخرى فإن هذه اللغة تصبح قادرة على توليد سائر بنيات اللغة وفق قوانين الاستعمال العادي لها»[8].

يمكن أن نجزم أنه لا سبيل إلى بناء المعاجم الآلية دون جعل الحاسوب قادرا على صياغة قواعد صورية بناء على المعطيات المخزنة لأن هذه القواعد هي التي ستساعده على استحضار مبدأ القياس أثناء عملية الإنتاج اللغوي وبالتالي سيستطيع توليد لغة عقلانية تكاد تماثل لغة المتكلم البشري، لأن المستوى المعجمي من أصعب المستويات أثناء الإنتاج اللغوي لما تعرفه اللغة العربية من تعقيدات دلالية فالكلمة تحمل معان متعددة وكل معنى يناسب سياقا محددا.

ولعل من مميزات المعجم الإلكتروني القدرة على استحضار الكمية الكبيرة والمعقدة من المعلومات، وبناء قاعدة بيانات معجمية للمفردات اللغوية بنوعيها البسيط والمركب، ثم فحص النص المعجمي ومراجعته، وإنشاء الرسائل انطلاقا من قواعد مضبوطة وصارمة، «بناء المعاجم الآلية يجب أن تكون متصفة بالسهولة والاطراد والوضوح للتوسع المستقبلي، كما يجب علينا وضع برنامج تحليلي لطبيعة المدخل المعجمي وما يمكن أن يحتويه من معلومات صوتية وصرفية ونحوية ودلالية»[9].

3.3 المستوى الصرفي:

تسعى الدراسات الصرفية إلى وصف القواعد التي تشكل البنية الداخلية للكلمات وطرق البناء من الناحية التشكيلية والتركيبية للقوالب اللغوية والصيغ والأوزان الصرفية وما يتصل بها من لواصق، وتتناول المفردات من حيث ترتيب أصواتها واشتقاقاتها وأجزائها وزياداتها، وتبحث في العناصر الصرفية أو الوحدات المورفيمية، التي من وظائفها الخطابية الربط بين أجزاء الجملة والكلمات، وهي ثلاثة أنواع: مورفيمات تدخل في مجالها العنصر الصوتي، مورفيمات لا تدخله، مورفيمات صورية خالية من العلامات الصرفية، و«يعد الصرف أحد فروع اللغة العربية المهمة، فعليه تبنى الكلمة، بل يعد الأساس الذي تقوم عليه

أصولها »[10] ولعل ما يوضح لنا أهمية علم الصرف هو النص الذي أورده ابن جني حينما ربط بيم علمي الصرف والنحو.

يقو ابن جني:«فالتصريف إنما هو لمعرفة أنفس الكلم الثابتة، والنحو إنما هو لمعرفة أحواله المتنقلة...وإذا كان ذلك كذلك فقد كان من الواجب على من أراد معرفة النحو أن يبدأ بمعرفة التصريف، لأن معرفة ذات الثابتة ينبغي أن يكون أصلا لمعرفة حاله المتنقلة، لأن أن هذا الضرب من العلم لما كلن عويصا صعبا، بدئ قبله بمعرفة النحو ثم جيئ به بعد، ليكون الارتباط في النحو موطنا للدخول فيه، ومعينا على معرفة أغراضه ومعانيه»[11]

إن علم الصرف يعتمد في مسائله على ما تقدمه له البحوث الصوتية من نتائج، وله دور كبير في الدراسات النحوية والمعجمية، من خلال تحديده أصول أبنية الكلمة بصيغها الأصلية والعارضة وما يطرأ عليها من تغيير معنوي في مدلولها، والتي لا يمكن فهمها دون دراسة للأصوات وخاصة عند الابتداء والوقف والتخفيف والتثقيل، لاسيما أن للنبر والتقسيم دورا في تكوين العلامة الصرفية[12]، فالمفهوم الصرفي يحدد بواسطة الجهاز الصوتي، والسياق، وحركات أوائل الكلمات ووسطها، والصيغ التي تعبر عن الوحدات الصرفية.

4.3 المستوى النحوي والتركيبي:

يعمل المستوى النحوي التركيبي على تنظيم الكلمات وترتيبها داخل الجمل والنصوص باستحضار ثلاثة مستويات أساسية تجمعها بهذا المستوى علاقة تأثير وتأثر، وهذه المستويات هي:

- المستوى الصرفي.
- المستوى المعجمي.
- المستوى الدلالي.

إن هذه المكونات الثلاثة يجب اعتمادها لبناء المعاجم الإلكترونية، فالمدقق الإملائي مثلا يقوم على الخوارزميات اللسانية في المستويين الصوتي والصرفي، حتى يتعرف الحاسوب على بنية الكلمة العربية من خلال القواعد التي يضعها اللسانيون، لا من خلال معجم الكلمات المزود به، «ورغم أن أغلب برامج معالجة النصوص العربية تذكر من بين خياراتها العادية التدقيق الإملائي والتدقيق النحوي إلا أنها لا تشتغل إلا بما هو موثوق سلفا على

الجهاز من المخزون المعجمي ولا تعتمد على محلل صرفي أو نحوي، ولهذا قد تخطئ كلمات أو جملا صحيحة لغويا، بسبب غياب الخبرة اللسانية الحاسوبية في هذه التطبيقات. كما أن برامج التشكيل الآلي لم تنجح لحد الساعة لعدم اعتمادها المحلل النحوي الذي أساسا لم يطور بعد.»[13]

وعلى ضوء هذه المستويات يعمل المستوى النحوي على تحديد ترتيب المفردات في الجمل بناء على ما تحمله من معان معجمية وما تؤديه من دلالات باستحضار المستوى الصرفي وما تحمله الكلمات من دلالات صرفية خاصة إذا كان الأمر يتعلق بعمل بعض المشتقات كاسم الفاعل واسم المفعول وغيرهما.

5.3 المستوى الدلالي:

تجدر الإشارة إلى أن المستوى الدلالي يلعب دورا مهما في الترجمة وتقديم المصطلح لكنه في غالبه لا يخرج عن حدود مساعدة المترجمين البشر له، كما أنه يساهم في عملية التوليد الآلي للكلام وهو ما يتطلب استحضار المكون الدلالي أثناء عملية الإنتاج حتى لا تُنتج جمل ونصوص صحيحة من الناحية التركيبية ولاحنة من الناحية الدلالية، ويقصد بالتوليد الآلي للغة «إنشاء نص بلغة طبيعية من نظام تمثيل إلى مثل قاعدة معرفة أو استمارة منطقية، وبعض الناس يعتبر التوليد الآلي للغة كمقابل لفهم اللغة الطبيعية، وفي نظم إنشاء التوليد الآلي للغة، يحتاج النظام إلى اتخاذ قرارات بشأن كيفية صياغة أحد المفاهيم. وأنجح التطبيقات للتوليد الآلي أنظمة تحويل البيانات إلى نصوص التي تقوم بإعداد خلاصات نصية للبيانات الرقمية وغير اللغوية حيث تمزج بين تحليل البيانات والتوليد الآلي للغة مثل النشرات الآلية لأحوال المناخ والاقتصاد والبورصة.»[14]

ومع زيادة توسع دائرة الحاجة لترجمة المعطيات من وإلى اللغة العربية ازدادت الحاجة إلى إسناد هذه المهمة للحاسوب وهو ما أدى إلى ظهور نوعين من الترجمة الآلية؛«الترجمة الأولى مسعفة بالحاسوب، وتعتمد على مخزون مشفر من قواعد البيانات: كلمات وتعبيرات وجمل مع مقابلاتها من اللغة الهدف، وتوكل إلى الحاسوب ترجمة أولية تقتصر على المعاني المعجمية للمفردات التي يتألف منها النص، لعدم مقدرته على ترجمة المعاني المجازية للمفردات التي تكتسبها في السياق التركيبي، مما يستدعي تدخل المترجم لإعادة تحرير النص.

الترجمة الثانية آلية ترتكز على المعرفة اللغوية الدقيقة في جميع المستويات اللغوية: الصرفية والنحوية والمعجمية والبرغماتية، ولا يوجد برنامج فعال للترجمة الآلية التامة من العربية وإليها، لتأخر البحث اللساني العام وافتقارها لأرضية عتادية لسانية صلبة، فالمترجم الآلي يحتاج إلى مدقق إملائي ونحوي، مولد صرفي ونحوي، وهذه لم يكتمل تطويرها بعد. وقد ظهرت الترجمة الآلية في الولايات المتحدة الأمريكية سنة 1947 على يد الباحث وارن ويفر الذي يعد أول من استخدم الحاسوب في الترجمة، ثم سارت قدما في باقي دول العالم، حيث تجاوزت الترجمة الآلية فيما بين اللغات الأوربية نسبة 95 بالمائة، في حين لم تخضع اللغة العربية للترجمة الآلية إلا بنسبة 50 بالمائة، رغم حاجة العربية لهذا النوع من الترجمة في تزايد مستمر نظرا للكم الهائل من المعارف والعلوم الأجنبية التي تستغرق مدة طويلة جدا لترجمتها بالطريقة العادية، وقد ساعدت الترجمة الآلية على ظهور ما يسمى بالبنوك الآلية للمصطلحات التي تسهل عمليات التعريب والترجمة، وأهم بنوك المعلومات المصطلحية في الوطن العربي: "باسم" بالمملكة العربية السعودية، "مكانز" للشركة العالمية للإلكترونيات بالقاهرة، "المعربي" بمعهد الدراسات والأبحاث للتعريب بالرباط.»[15]

إن نجاح عملية الترجمة الآلية رهين بتعالق كل مستويات اللغة بالمستوى الدلالي خاصة وأن هذا الأخير لا يعمل بمعزل عن المستوى المعجمي الذي سبق وأن وقفنا عنده فوجدناه هو بدوره يعرف عدة تعقيدات لدى الآلة مقارنة بالإنسان ،فعمليتي إنتاج وإدراك الوحدات المعجمية عمليتان سريعتان للغاية، لأن « الإنسان يمتلك قدرة هائلة في معالجة كم هائل من الوحدات المعجمية أثناء عملتي الإنتاج والتلقي وفي هذا تأكيد أن الفرد يخزن المعلومات المعجمية في معجمه الذهني ويلج إليها أو يبني بعض الأجزاء الأخرى انطلاقا من آليات محددة».[16] وهذا ما يغيب لدى الآلة أثناء عملية الترجمة الأمر الذي يفرض علينا باستمرار الوقوف عند النصوص المترجمة ومراجعتها.

4. تحديات تطبيقات معالجة اللغة العربية و آفاقها:

إن صعوبات المعالجة الحاسوبية لمستويات اللغة العربية تتمثل بشكل أساسي في نقص الخبرة اللسانية الحاسوبية ونقص في فهم الحاسوبي لمتطلبات اللغوي وانعدام التعاون بينهما، حيث «يرجع فشل البرمجيات التطبيقية العربية إلى اعتماد الخبرة الحاسوبية وإغفال الخبرة اللسانية الصورية، فقد اعتمد المبرمجون على بعض اللغويين غير القادرين على فهم متطلبات الحاسوب، فزودوهم بمعلومات أكاديمية صحيحة إلا أنها

غير قادرة على الاستجابة لحاجة المبرمجين، فالغرب لم يشرعوا في وضع برامجهم التعليمية إلا بعد أن اكتمل لديهم الوصف اللساني الحاسوبي للغتهم، أما نحن العرب فلم يكتمل ذلك الوصف، وجاءت برامجنا التعليمية محصورة في الخبرة التربوية والخبرة الحاسوبية، أما الخبرة اللسانية التي تشكل عصبها المركزي فلم تؤخذ في الحسبان، مما أدى إلى عدم نجاعتها.»[17] خاصة وأن مدخل حوسبة اللسان العربي يكمن في فهم منطقه الداخلي وصورنة علاقاته وتحويلها إلى رموز يسهل تلقينها للحاسوب بشكل يجعله يستحضرها أثناء إنتاج أو استقبال الفعل الكلامي لتحاكي قدرته قدرة المتكلم المستمع المثالي حسب تصور اللسانيات التوليدية.

وعلى الرغم مما وصلت إليه تطبيقات برامج حوسبة اللغة العربية ما زالت تعترضها بعض التحديات التي تحول دون رقيها إلى مصاف تطبيقات ألسن أخرى، ويمكن إجمال هذه التحديات في:

1.4 اللبس الدلالي:

يختلف الفرق في اللبس بين القارئ البشري والقارئ الآلي في برنامج ترجمة آلية أو تلخيص آلي أو برنامج تشكيل آلي، فالقارئ البشري لا تشكل له الترجمة أو التلخيص أو التشكيل لبسا لقدرته على فهمها بسبب ثقافته وإلمامه بالسياق واستخدامه للمنطق، أما القارئ الآلي فسيعاني من اللبس في الوحدات اللغوية المتعددة المعاني، لذلك ينبغي «تحويل القيود اللغوية – النحوية والدلالية- إلى لوغاريتمات تساعده على اختيار المعنى المقصود»[18]

وإشكالية اللبس في اللغة العربية مشكلة ليست سهلة أو هينة، وتشكل تحديا صعبا يواجه المعالجة الآلية للغة العربية، ولحل مشكلة اللبس يجب «حصر جميع الوحدات اللغوية(الأفعال وحروف الجر والأسماء والمصادر) اللبسية بجميع صور اللبس ووضع التفسير المناسب لكل سياق، ثم العمل على القاعدة التي تحدد القيود المعجمية والنحوية والتركيبية والدلالية اللازمة لفك اللبس»[19]

ولا يكتمل النظر في إشكالات اللسانيات الحاسوبية العربية دون الوقوف عند قضية غياب التشكيل وما تطرحه من إشكالات، فالكلمة العربية حينما تجرد من التشكيل لا تتعدد حمولاتها الدلالية فقط وإنما حمولاتها الصرفية والنحوية أيضا، هذا بالنسبة للحاسوب فقط أما متكلم اللغة العربية فبقدرته اللغوية الذهنية قادر على تحديد الدلالة الصرفية والنحوية حتى لو تجردت الكلمة من الحركات.

ويمكن أن يختلط الأمر على الحاسوب فيلتبس عنده الأمر في جملة ما لغياب التشكيل فكلمة "وجد" تحتمل معنيين الأول دال على معنى "عثر" وذلك إذا اعتبرنا الواو جزءا من الكلمة والثاني دال على الاجتهاد وذلك حينما نعتبر الواو حرفا عطف، كما أن كلمة سر يحتمل أن تكون فعل أمر أو اسم ولا يظهر ذلك إلا بالحركة الإعرابية، والأمر نفسه في كلمة "أسر" التي تحتمل أربعة أوجه:

- أسَرّ الحديث
- أَسَرَ الجندي
- أُسَرٌ جمع أسرة
- أُسُر ما يقيد به الأسير من حبل أو قيد.

2.4 اللبس التركيبي والنحوي:

في المستوى التركيبي النحوي يمكن لجملة ما أن تدل على أكثر من معنى ويمكن تفسيرها بأوجه مختلفة، ففي جملة "عصفور البستان الجميل" يمكن أن تحيل الصفة "الجميل" على العصفور والبستان جميعا، والحال في مثل شاهدت الأستاذ متوترا فالالتباس هنا تحديد من المتوتر أهو الفاعل أو المفعول والعطف بين عدة وحدات لغوية كما في مثل سلمت على الرجال والنساء والأرامل، فمن الأرامل وجملة ضربت الرجل ضاحكا فمن الضاحك هل الرجل أم الفاعل في الجملة، ويمكن أن نضيف إلى هذه الإشكالية التي يعاني منها المستوى النحوي التركيبي إشكالية تغير معنى الأفعال تبعا للوحدات اللغوية المجاورة فرغب في ليست هي رغب عن.

4.3 تداخل مستويات اللسان العربي وتفاعلها:

يظهر ذلك من خلال التماسك بين النحو والصرف والتداخل الشديد بينهما وشدة الصلة بين أشكال الصيغ الصرفية ومعانيها وهذا التماسك سلاح ذو حدين على قول نبيل علي، «فهو من جانب يزيد من صعوبة النظم الآلية حيث يصعب فصل المعالجات الآلية للنحو الآلي عن المعالجات الآلية للصرف العربي والمعجم العربي، ويتطلب ذلك حاسبات ذات إمكانات أكبر وقدرة أكبر على معالجة اللغة العربية آليا، ومن جانب آخر يعمل هذا التماسك على إجلاء اللبس الناجم عن غياب التشكيل حيث توفر قرائن لغوية تكشف اللبس.»[20]

جلي إذن أن تماسك مستويات اللسان العربي له القدرة على إمكانية فك اللبس الذي يمكن أن يواجه معالجة اللغة العربية آليا، وذلك بمعرفة نقاط الاتصال والانفصال بين هذه المستويات، فالعلاقة الجدلية التي تكمن بينها ظاهرة لكن القواعد المنطقية المتحكمة في هذه العلاقة للأسف لم يُتوصل بعد إلى استنباطها بشكل تكون الآلة قادرة على استقبالها وتحويلها إلى وجود لغوي مادي، خاصة وأن «المستوى النحوي والصرفي العربيان توأمان يصعب الفصل بينهما مهما زادت دقة المشرط التحليلي. وقضية فك الاشتباك الصرفي – النحوي تأتي على قائمة الأولويات في فهم أداة اللغة العربية بأكملها، وتعد المنطق الأساسي لمعالجة النحو العربي آليا. ومن الضروري أن نستخلص من زخم العلاقات المتشابكة قائمة محددة بالسمات الصرفية النحوية للغة العربية في صورة هيكلة سليمة تعكس العلاقات البنيوية التي تربط بين هذه السمات.

إن هذه السمات لا بد أن تأتي مكتملة لتلبي مطالب كافة النظم اللغوية، وكذا تلك الخاصة بمعالجتها آليا، وخلاصة ذلك فإن السمات الصرفية النحوية هي إحدى الركائز الرئيسة التي لابد منها لصياغة النحو العربي، في صورة رسمية دقيقة وبما يهيئه للمعالجة الآلية»[21]، لأن حضور هذين المستويين عند المتكلم البشري مخالف تماما لحضورهما عند المتكلم الآلي، لأن الأول مكتسب لخوارزميات هذا اللسان فطريا، أما الثاني فهو مزود بها ولا يحسن التعامل معها إلا إذا سهلنا له سبل ذلك بتحديد العلاقات المتحكمة في عملها.

4.4 غياب البعد السياقي في الفعل اللغوي الحاسوبي:

إن حل لغز الذهن البشري يحتاج إلى توفير آليات وبرامج لغوية تستحضر خصوصية كل لسان، لذلك كان الهدف الأسمى للنظم الآلية للتحليل اللغوي هو الوصول إلى نظام أوتوماتيكي لفك السياق اللغوي في صورته المنطوقة والمكتوبة، ولو أخذنا على سبيل المثال نظام الفهرسة الإحصائية الآلية وجدناها تعاني من مشاكل مردها إلى مقتضيات السياق وموضوع الحديث كالترادف والتضاد والاشتراك اللفظي، وعليه يكون صلب استرجاع المعلومات يكمن في معنى السياق قبل معنى الكلمة بلا شك.

ومن المقرر أن أمر الترجمة الآلية علم تقني ينبني على دمج المعرفة البشرية بالنظم الحاسوبية من خلال البرمجة الرقمية، لكن هذا الدمج وعلى الرغم مما قدمه من إسهامات وخدمات إلا أنه لم يخل من مجموعة من العيوب «ويرجع أول هذه العيوب إلى اللغة العربية نفسها، فإنه منذ سنوات تجد أن في الأسواق برمجيات عربية للترجمة الآلية إضافة

للقواميس الإلكترونية وبرمجيات التحليل النحوي والصرفي والتدقيق الإملائي للنصوص العربية. وبدخول الإنترنيت في الدول العربية وما يوفره من فرصة عظيمة للحصول على المعلومات، أصبح الباب مفتوحا على مصراعيه للاستفادة من الترجمة الآلية العربية في ترجمة تلك المعلومات والخزين الهائل من مصادر العلوم والمعرفة وما يظهر من جديد إضافة إلى مواقع الويب والصحافة والأخبار وغيرها مما لا حصر له من مصادر المعلومات المكتوبة فمن هنا ظهرت بعض عيوب الترجمة الآلية»[22]، والقول بعيوب اللغة العربية لا يتصل البتة بمفهوم العيب في ذاته وما يحمله من دلالة قدحية، وإنما يرتبط بما ينتج من خلل في الترجمة الآلية، فالنقص هنا إنما يتصل بنقص في وظيفة الترجمة الآلية لا في اللغة العربية، خاصة وأن الترجمة البشرية لا ينتج عنها أي نقص أو تحريف للمعنى نظرا لاستحضار المترجم سياق اللفظة في الجملة وكذلك المستوى التداولي وهذا الأمر يغيب في الترجمة الآلية نظرا لغياب قدرة ذهنية تستحضر كل مستويات الكلام من صرف ودلالة وتركيب ومعجم وسياق، فلو أخذنا لفظة "عين" على سبيل المثال وجدنا حمولتها الدلالية تختلف من سياق لآخر فقد تدل« على عضو به يبصر وينظر[...]قال الخليل العين الناظرة لكل ذي بصر[...]ويقال العين الجارية النابعة من عيون الماء[...] ومن الباب العين السحاب ما جاء من ناحية القبلة[...].»[23]، فلفظة "عين" إذا ما وردت في نص نريد ترجمته للغة العربية وجدت الآلة نفسها تائهة بين عدة حمولات دلالية وهو ما يستلزم تزويد الآلة بمعجم حاسوبي مبني وفق خوارزميات تميز كل لفظة عن الأخرى، بل وتميز الحمولات الدلالة التي تحملها المفردة الواحدة وهو ما يدخل في قضية المشترك اللفظي.

لذلك فلا غرابة أن تكون جهود البحث العربي في هذا المجال متواضعة إذا ما قورنت بالجهود الغربية فيه، فقد «تجاوزت البحوث المقدمة باللغة الإنجليزية خدمة للعربية ثلاثين ألف بحث، بينما لا تتجاوز الدراسات باللغة العربية المئات، ومعظم الشركات المفعلة لهذه البحوث هي شركات عالمية محركة للشركات المستخدمة دوليا مثل مايكروسوفت وأبل وجوجل وسامسونج، غير هذه التطبيقات جاءت وفق حاجاتها الملحة لزيادة رأس المال، فهي لم تعن بالجوانب الاجتماعية والفكرية للعربية؛ لذلك ظلت متأخرة في خدمتها الحاسوبية عن اللغات الأخرى، ولا ننسى احتكار هذه الشركات للتطبيقات المنفذة من قبلها، وأن ما قامت به من تقنيات حوسبة اللغة العربية إنما هي معالجة جزئية لبعض جوانبها، ولم تكن تحمل مشروعا عاما له إطاره اللغوي والفكري»[24]

وهو ما يؤكد أن مستقبل حوسبة اللسان العربي في أمس الحاجة إلى نظرة أشمل تتجاوز حدود الخطاب اللغوي الراهن المبني على أسس وصفية للفروع لا الأصول لتصل إلى ما يحكم هذا الخطاب من قواعد«سواء من قبل اللغويين وأهل المعاجم وأئمة المجامع، أو من قبل نقاد الأدب والتربويين والإعلاميين. لقد باتت إشكالية اللغة العربية، من المحورية والشمولية والتعقيد، بحيث يستحيل تناولها انطلاقا من منظور التخصص الضيق، أو النظرة الاجتماعية القاصرة، ومن الخطورة والأهمية، بحيث يصعب تناولها، من دون استراتيجية واضحة للإصلاح اللغوي الشامل؛ وذلك في إطار خطة قومية أكثر شمولا، لإعداد مجتمعاتنا العربية لدخول عصر المعلومات»[25]، وهو ما يفرض على الباحثين في مختلف المجالات المرتبطة باللسان العربي العمل بشكل جماعي لتحقيق هذه الغاية.

5. خاتمة:

إذا كانت اللغة العربية تختص بالكفاية اللسانية شأنها شأن باقي الألسن الأخرى فإن ما تتميز من خصائص جبرية منطقية تجعل من النظم الحاسوبية قادرة على معالجتها آليا، لكن شريطة توفر الشروط الموضوعية التي تيسر ذلك، وفي ظل غياب فهمها في شموليتها انطلاقا من جذورها الصورية لا فروعها المحققة تبقى الأنظمة المبتكرة هي بدورها عاجزة عن فهمها، فمن غير الممكن بل من المستحيل أن تنجح الآلة في محاكاة الفعل الكلامي في كل مستوياته بنسبة مائة بالمائة والمتكلم المستمع المثالي لهذه اللغة غير قادر على الكشف عن القواعد التي تمكنه من إنتاج هذا الفعل، لذلك يمكن أن نؤكد أن الخبرة اللسانية هي المدخل الأول لنجاح نظم حوسبة اللغة العربية خاصة وأن هذه اللغة مازالت تعرف العديد من الإشكالات العالقة منها ما يرتبط بالمستوى الصرفي كاشتقاق المصدر من الفعل الثلاثي وانتقاء الكلمات لجمع تكسير دون غيره وتداخل الصيغ الصرفية، ومنها ما يرتبط بالمستوى الدلالي، فحينا ننجح في إيجاد حلول لسانية منطقية لهذه الإشكالات وغيرها سنتمكن لا محالة من ولوج عالم حوسبة اللغة من أوسع أبوابه إذ سيسهل علينا نقل قواعد وخوارزميات هذه اللغة للمبرمجين المشتغلين في هذا المجال.

مراجع البحث وإحالاته:

1- نبيل علي، اللغة والحاسوب، ص 3.

2- عبد العالي العامري، اللغة ونظرية الذهن، مجلة اللسانيات العربية، تصدر عن مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، العدد 6، يناير 2018. ، ص 9

3- أبو الفتح عثمان بن جني ، الخصائص، ج 1، ص 33

4- محمد عرباوي، تأثير مرجعيات الترجمة والحوسبة في محاكاة وتمثيل اللغة العربية، ص 23.

5- أبو الحجاج محمد بشير، المعالجة الآلية للغة العربية، جهود الحاضر وتحديات المستقبل، ص 19.

6- أبو الحجاج محمد بشير، المعالجة الآلية للغة العربية، جهود الحاضر وتحديات المستقبل، ص 19.

انظر 7 - شركة الهندسة لتطوير النظم الرقمية في مجال حوسبة اللغة العربية https://www.rdi-eg.com/ موقع

8- محمد عرباوي، تأثير مرجعيات الحوسبة في محاكاة وتمثيل اللغة العربية، ص 19

9- باية سهام، اللسانيات الحاسوبية والمعجمية العربية، مجلة لغة – كلام، مجلد 03، عدد 2، 2017، ص 235.

10- محمد التاقي بمشاركة عايدة بنت سعيد البصلة، مقاربة صرافية مبتكرة لتشكيل الجذور الثلاثية المعتلة في اللسان العربي، منشورات كلية الآداب والعلوم الإنسانية بالرباط، سلسلة بحوث ودراسات رقم 92، مطبعة الكرامة، ط 1، 2021م، ص 6.

11- ابن جني، المنصف، ج 1، ص 4-5.

12- محمد عرباوي، تأثير مرجعيات الحوسبة في محاكاة وتمثيل اللغة العربية، ص 22.

13 - محمد عرباوي، تأثير مرجعيات الحوسبة في محاكاة وتمثيل اللغة العربية، ص 21.

14- أبو الحجاج محمد بشير، المعالجة الآلية للغة العربية، جهود الحاضر وتحديات المستقبل، ص 19.

15 - محمد عرباوي، تأثير مرجعيات الحوسبة في محاكاة وتمثيل اللغة العربية، ص 24.

16- لخزاز عبد الإله، الهندسة الداخلية للمعجم الذهني وتنظيم المعرفة المعجمية في ضوء اللسانيات العصبية، مجلة الميادين للدراسات في العلوم الإنسانية، المجلد الثاني، العدد الأول، ص 60

17 - محمد عرباوي، تأثير مرجعيات الحوسبة في محاكاة وتمثيل اللغة العربية، ص 25.

18- أبو الحجاج محمد بشير المعالجة الآلية للغة العربية، جهود الحاضر وتحديات المستقبل، ص 23.

19- أبو الحجاج محمد بشير المعالجة الآلية للغة العربية، جهود الحاضر وتحديات المستقبل، ص 23.

20- أبو الحجاج محمد بشير المعالجة الآلية للغة العربية، جهود الحاضر وتحديات المستقبل، ص 21.

21- نبيل علي ، اللغة العربية والحاسوب، تعريب، 1988م، ص 349.

22- عايض محمد الأسمري، الترجمة الآلية من منظور اللسانيات الحاسوبية، المجلة العربية للعلوم ونشر الأبحاث، العدد 3، المجلد 4، 2018،.ص 7.

23- أبو الحسن أحمد بن فارس بن زكرياء ، مقاييس اللغة، دار الفكر، لبنان، ط 1، 2011 مادة "عين" ص 726-727.

24- محمد السعودي ، اللغة العربية والحوسبة، مجلة أفكار، المحتوى الرقمي العربي على الإنترنيت، ص 29.

25- نبيل علي الثقافة العربية وعصر المعلومات: رؤية لمستقبل الخطاب الثقافي العربي ، ص 235.

قائمة مصادر البحث:

1. أبو الحجاج محمد بشير ، المعالجة الآلية للغة العربية، جهود الحاضر وتحديات المستقبل، مجلة لغة العصر، العدد 92، أغسطس 2008م.
2. أبو الحسن أحمد بن فارس بن زكرياء، مقاييس اللغة، دار الفكر، لبنان، ط 1، 2011م.
3. أبو الفتح عثمان بن جني ، المنصف، إدارة أحياء التراث القديم، الطبعة 1، 19954م
4. أبو الفتح عثمان بن جني، الخصائص، تحقيق عبد الحميد هنداوي، دار الكتب العلمية، بيروت – لبنان، الطبعة الرابعة، 2013م.
5. باية سهام، اللسانيات الحاسوبية والمعجمية العربية، مجلة لغة – كلام، مجلد 03، عدد 2، 2017.
6. عبد العالي العامري ، اللغة ونظرية الذهن، مجلة اللسانيات العربية، تصدر عن مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، العدد 6، يناير 2018م.
7. لخزاز عبد الإله، الهندسة الداخلية للمعجم الذهني وتنظيم المعرفة المعجمية في ضوء اللسانيات العصبية، مجلة الميادين للدراسات في العلوم الإنسانية، المجلد الثاني، العدد الأول.
8. محمد الأسمري، الترجمة الآلية من منظور اللسانيات الحاسوبية، المجلة العربية للعلوم ونشر الأبحاث، العدد 3، المجلد 4، 2018.
9. محمد التاقي بمشاركة عايدة بنت سعيد البصلة، مقاربة صرافية مبتكرة لتشكيل الجذور الثلاثية المعتلة في اللسان العربي، منشورات كلية الآداب والعلوم الإنسانية بالرباط، سلسلة بحوث ودراسات رقم 92، مطبعة الكرامة، ط 1، 2021م.
10. محمد السعودي، اللغة العربية والحوسبة، مجلة أفكار، المحتوى الرقمي العربي على الإنترنيت.
11. محمد عرباوي، تأثير مرجعيات الترجمة والحوسبة في محاكاة وتمثيل اللغة العربية، مجلة حوليات الأدب واللغات، جامعة محمد بوضياف –المسيلة، 2013م.
12. نبيل علي، الثقافة العربية وعصر المعلومات: رؤية لمستقبل الخطاب الثقافي العربي، علم المعرفة. يناير 2001م.
13. نبيل علي، اللغة العربية والحاسوب، تعريب، 1988م.