المعالجة الآلية للأطلس اللساني الرقمي الجزائري

- مقدمات منهجية في التصميم والإنجاز-

The automatic processing of the Algerian digital linguistic atlas Systematic Introductions to Design and Delivery

د. إبراهيم براهمي *
جامعة 8 ماي 1945 . قالمة(الجزائر)

Dr.brahim brahmi
University 8 May 1945 Guelma(Algeria)
brahimi12@hotmail.com

تاريخ استلام المقال: 2021/05/03	تاريخ القبول: 2021/09/14	تاريخ النشر: 2021/09/30

ملخص

موضوع الدراسة بيان دور الذي تضطلع به المعالجة الآلية في صناعة أطلس لساني جزائري رقمي؛ وإبراز كيفية استثمارها في تمثيل المعطيات اللغوية المرتبطة بهذا الأطلس؛ من نحو: تصميم الخرائط اللسانية وإنجازها. ومعالجة الظواهر اللغوية المنطوقة والمكتوبة في حيزها الجغرافي المحدد. وتحاول الدراسة استكشاف مجالات الاستفادة من التكنولوجيات الحديثة وتمظهراتها الرقمية في وضع برمجيات. وبناء أرضية بيانات وطنية لهذا الأطلس اللساني. تخزن فيها كل ما يرتبط بهذا لمشروع الوطني. بحيث تتيح لكل باحث لغوي الولوج إليها بطريقة سلسة حركية وتفاعلية؛ لتطرح عديد الأسئلة من نحو: كيف تسهم الرقمنة في صناعة هذا الأطلس اللساني وتوطين محتوياته؟ ما هي الآليات العملية في حوسبة هذا الأطلس اللساني؟ وما طبيعة التحديات التي تعترض عملية حوسبة المعطيات اللغوية للأطلس اللساني الجزائري.

الكلمات المفاتيح: المعالجة الآلية؛ الأطلس؛ اللساني؛ الجزائر؛ الرقمنة؛ اللهجات

Abstract

The topic of the intervention is an explanation of the role that computational linguistics plays in the making of a digital Algerian linguistic atlas. Show how it is invested in the

* **المؤلف المرسل:** إبراهيم براهمي **البريد الإلكتروني:** brahimi12@hotmail.com

representation of the linguistic data associated with this atlas. In terms of: the design and implementation of linguistic maps, and the automatic treatment of spoken and written linguistic phenomena in their specific geographical space according to the levels of the linguistic system.

The intervention tries to explore aspects of the relationship between language, informatics and digitization, and areas of utilization of modern technologies and their digital manifestations in developing software, and building a national data platform for this linguistic atlas, in which everything related to this national project is stored, and that each linguistic researcher can access it in a smooth, dynamic and interactive way, and it helps him In studying linguistic codes, editing, classifying and analyzing texts; To ask many questions from about: How does digitalization contribute to the making of this linguistic atlas and the localization of its contents? What are the practical mechanisms in the computing of this linguistic atlas? And what are the nature of the challenges that could be encountered in the process of computing the linguistic data of the Algerian linguistic atlas?

Keywords: automatic processing; atlas; linguistics; Algeria; digitization; dialects.

1. مقدمة

أحدث تطور تكنولوجيا المعلومات والاتصالات المتسارع في مطلع هذه الألفية بناء تصورات جديدة في سيرورة اكتساب المعارف والعلوم. حتى غدا من أوضح سمات هذا العصر الرقمنة والمعلوماتية. وحوسبة مختلف أنساق المعرفة؛ ولعل من أبرز المظاهر المتجلية لهذا التطور شيوع توظيف الحاسوب وبرمجياته المتنوعة (نظرا للإمكانات الهائلة التي يمتلكها الحاسوب في استقبال البيانات. والمعلومات ومعالجتها وتخزينها. وتقديمها بسرعة ودقة؛ فقد شاع استخدامه في مختلف مجالات الحياة في الطب. والهندسة. والزراعة. والصناعة. والفضاء. والبحوث العلمية)(مختار عبد الخالق عبد الله.2007: 112)؛ فأمست البرمجيات الحاسوبية وتكنولوجيا المعلومات أداة مؤثرة تسهم بفاعلية في تشكيل الوعي الوجداني للمجتمعات الإنسانية المعاصرة؛ وتوجيه الرأي العام العالمي. ومن جانب آخر لم يعد توظيف هذه البرمجيات الحاسوبية مقتصرا على علوم المادة البحتة؛ كما كان الشأن عليه في بادئ الأمر. بل أضحت العلوم الإنسانية والاجتماعية المجال الرحب لتطبيقاتها المتعددة.

وفي هذا الخضم أسست اللسانيات لعلاقات بينية (Interdisciplinaire) أمبريكية مع البرمجيات الحاسوبية وتكنولوجيا المعلومات؛ منذ أن ظهر الحاسوب في أشكاله الأولى في أربعينيات القرن الماضي. وواكب ذلك انفتاح النسق اللساني. وظهور مختلف النظريات اللغوية ونماذجها

 E-ISSN - 2773-2886 / ISSN - 2773-2894

اللسانية التنظيرية والتطبيقية المتنوعة. (وقد أثمر هذا التفاعل العلمي بين اللغة والفروع العلمية المختلفة عدة نماذج لغوية أسهم في ظهورها بصور مباشرة وغير مباشرة. التوسع في استخدام الكمبيوتر في المجال اللغوي وما تتطلبه من تمثيل اللغة بصورة منضبطة تلبية لمطالب المعالجة الآلية)(نبيل علي. نادية حجازي.2005: 323).

إن الحديث عن المعالجة الآلية الحاسوبية للأطلس اللساني الجزائري. والدعوة إلى تأسيس أطلس لساني رقميٍّ؛ يعني في واقع الحال وضع الاستعمال اللغوي الجزائري المكتوب والمنطوق في صورة خرائط تبيِّن التَّوزيع الجغرافيّ للأنماط اللغويَّة- فصيحها وعاميها- واستخداماتها؛ في شكل تطبيقات رقمية وبرامج الكترونية متنوعة؛ تتيح لاستخدامه والاطلاع عليه على نطاق واسع. ولذلك فإن ما ترمي إليه هذه الورقة هو الإجابة عن أوجه استثمار الحاسوبيات في إنجاز الأطلس اللساني الجزائري. وتحويله من كتل معرفية ثابتة غير متاحة للجميع. إلى كتل معرفية رقمية ذات جودة عالية. وحركية فاعلة في مجتمع المعرفة الإنساني. غير إن التساؤل يظل عالقا عن طبيعة هذه المعالجة الآلية وأنماطها ومستوياتها؟.

ومن أجل استكشاف هذه الجوانب جاءت هذه الدراسة في مقدمة وأربعة مباحث:

- أوضحت في المبحث الأول ماهية المعالجة الآلية من حيث الاصطلاح والوظيفة.

- وفي المبحث الثاني عرضت للأطلس اللساني الجزائري من الورقي إلى الرقمي.

- وفي المبحث الثالث لدور المعالجة الآلية ومنهجيتها في صناعة الأطلس اللساني الجزائري الرقمي.

- وفي المبحث الرابع لمداخل استثمار المعالجة الآلية في صناعة أطلس لساني جزائري رقمي؛

- ثم خاتمة متضمنة لأهم النتائج المتوصل إليها.

2. مفهوم المعالجة الآلية

يعد اصطلاح المعالجة الآلية من أكثر المفاهيم استعمالا وشيوعا في الدرس اللساني المعاصر. ويرتبط في استخدامه محوريا باللغات الطبيعية Natural Language Processing ضمن هذا الحقل المعرفي الجديد الذي يعرف بـ"اللسانيات الحاسوبيةComputational Linguistics". ولإدراك دلالات المفهوم وجب تحليله في بنيته اللغوية الظاهرة؛ التي تبدو في صورته المبسطة؛ اسما مركبا تركيبا إضافيا؛ يجمع في تكوينه بين صفة وموصوف؛تربط بين جانبين وظيفيين وهما اصطلاحا المعالجة والآلية (التقانة) :

- المعالجة ويراد بها (التطبيقات الآلية على مجموعة نصوص من اللغة. وذلك بتحويلها وبتغييرها. وإبداع شيء جديد اعتمادا عليها. ويتم ذلك باستعمال تقنيات وأدوات من اللسانيات والإعلام الآلي والنمذجة (Modélisation). ومن الواجب التفرقة بين وصف المعارف التي هي من وظيفة اللسانيات. والتعبير عن هذه المعارف في نماذج باستخدام تقنيات. واستراتيجيات فعالة من علم الحاسوب؛ وهي وظيفة علم اللغة الحاسوبي)(بن عربية. راضية.2017: 38).

- وفي تحديد اصطلاحي آخر؛ (فإن المقصود بالمعالجة في مجال المعلومات؛ هو إجراء سلسلة من الإجراءات. أو العمليات على معلومات محددة خاصة بموضوع ما بغرض تحقيق نتائج معينة يحددها تخطيط للوصول للحل) (السالمي. علاء عبد الرازق.2002: 73).

- أما اصطلاح الآلية فيراد بها تلك العمليات التي تنجز عن طريق الآلة والتي تقابلها العمليات التي ينجزها الإنسان؛ والحاسوب هو الآلة التي تستعمل في معالجة اللغة. واخترعت لإجراء العمليات الحسابية. ووجب تطويرها لمعالجة المعلومات ذات الصلة بالطبيعة اللسانية؛ فالمعالجة الآلية: هي تتابع حركات حسابية تقوم بها الآلة. وفق تسلسل زمني؛ وهذا يعني أن برنامج المعالجة الآلية (Traitement Automatique des Langues) يمكن أن يكون كليا(Total) أو جزئيا(Contraites)؛ بحيث أن:

E-ISSN - 2773-2886 / ISSN - 2773-2894

1. الكلي: يقوم فيه الحاسوب بكل شيء

2. الجزئي: يتدخل الإنسان في بعض المراحل(بن شاشة. فارس.2008: 38).

وإذا نظرنا إلى أنماط المعالجة الآلية فلمختلف الحاسبات أنماط من المعالجة نذكر منها:

- المعالجة المتزامنة (Concurrent Processing)

- المعالجة المتداخلة (Overlapped Processing)

- المعالجة الدينامكية (Dynmaic Processing)

- المهام المتعدد (Multitasking)

- البرمجة المتعددة (Mutiprogramming)

- المعالجة المتعددة (Multiprocessing) (السالمي. علاء عبد الرازق.2002: 74).

إن من أهم ما يمكن أن يستخلص من هذه التحديدات الجزئية للمعالجة الآلية؛ ذلك الارتباط الوظيفي بين المصطلح واللسانيات الحاسوبية. والذكاء الاصطناعي (Artificial Intelligence) من جانب آخر.

إن المعالجة الآلية للغة بهذه الصورة؛ هي عملية ابتكار أنظمة و"برامج معلوماتية". و"محطات عمل" تحلل النصوص المكتوبة والمنطوقة آلياً؛ أي دون تدخل الإنسان. وذلك من أجل مساعدة مستعمل الحاسوب على حل بعض الأمور. والمسائل المتعلقة باللغة؛ مثل: برامج الترجمة الآلية. التصحيح الإملائي والقواعدي. والخلاصة الآلية. وتحليل الخطاب آلياً. والإحصاء الآلي للمفردات. وبرامج إنتاج المصطلحات(غسان. مراد.2019).

وتمر المعالجة الآلية للغة بمراحل عدة. لكل مرحلة مشاكلها وحلولها: كالتجزئة الآلية للنصوص. ومرحلة التحليل الصرفي. والنحوي. والدلالي... الخ. هذه المراحل ببنيتها تشبه الطبقات. وكل طبقة تعتمد على سابقتها؛ هذا بالنسبة لما يسمى المعالجة الكلاسيكية للغة. إذ يوجد حالياً طرق أخرى تعتمد على المدونات. إن المعالجة الآلية للغة تعتمد بشكل أساسي على الظواهر والقواعد اللغوية. فهي تتطلب دراسة ألسنية معمقة لكل المراحل المذكورة آنفاً. إذ لا يمكن أن تتم هذه العملية إلا بالاعتماد

على النظريات اللغوية المختلفة التي بدورها تساعد على رفع الالتباس وتحديد المعاني لكل جزء نصي حسب السياق اللغوي. والمفردي والصرفي والنحوي والدلالي(العناتي. وليد. 2012: 23).

وفي المجمل فإن المعالجة الآلية للغة هي تمثيل حي للسانيات الحاسوبية. هذا العلم المعرفي الذي يتداخل مع حقل الذكاء الاصطناعي الهادف إلى وضع نماذج حاسوبية للإدراك الإنساني؛ ومنتهى غايتها أن تبلغ بالحاسوب مَبْلَغَ الكفاية اللغوية التي يمتلكها الإنسان حين يستقبل اللغة وينتجها فيصبح قادرا على تركيب اللغة وتحليلها (غسان. مراد.2019).

3 . الأطلس اللساني الجزائري من الورقي إلى الرقمي

يشير مفهوم الأطلس اللساني(l'Atlas linguistique) في تحديداته اللسانية المعاصرة إلى (مجموعة من الخرائط الجغرافية التي توضح تنوع اللهجات لأشكال لغوية معينة. وغالبًا ما تظهر في صورة وحدات معجمية. ويمكن أن يشير المصطلح أيضًا إلى مجموعة من الخرائط التي توضح توزيع اللغات واللهجات لمنطقة جغرافية معينة.) (Georges Mounin, 2004 :44).

في تحديد اصطلاحي لغوي عربي معاصر يشير إلى كونه (مجموعة من الخرائط تبيِّن التَّوزيع الجغرافيّ للأنماط اللغويَّة واستخداماتها)(أحمد مختار عمر. 2008 : 101).

إن هذه التحديدات الاصطلاحية تشير إلى أربعة أمور:

أولها. التعريف الذي يشير إلى مجموعة خرائط؛ هذا الاستخدام لصيغة الجمع بدل الإفراد؛ هو مما يدل على أنه لا يوجد نمط واحد للخريطة بل هي في صورتها العملية مجموعة متنوعة ومتعددة في أشكالها وأهدافها.

وثانيها. أنه يحمل معنى التَّوزيع الجغرافيّ للغة؛ دلالة الانتشار للظواهر اللغوية وشيوعها في حيز مكاني معين.

وثالثها. أنه يوظف تركيب الأنماط اللغويَّة؛ والمراد بشكل أوضح اللغات أو اللهجات المتفرعة عنها؛ وما يرتبط بهما من ظواهر لسانية تشيع في ذلك الحيز الجغرافي.

ورابعها. أنه يشير إلى تركيب "استخدامات هذه الأنماط اللغويَّة"؛ والمراد الاستعمال اللغوي الذي قد يختلف عند أبناء الوطن الواحد من منطقة إلى أخرى ومن مكان إلى آخر.

E-ISSN - 2773-2886 / ISSN - 2773-2894

وتكتسي الخرائط اللغوية أهمية بالغة في دراسة اللغة وتتبعها ميدانيا. فمن شأنها أن (تطلعنا على الاختلافات الصوتية بين المناطق المختلفة؛ فقوم يجهرون أصواتا. وقوم يهمسونها. وطائفة تنطق الفتحة صريحة. وأخرى تنطقها ممالة. ولهجة تنبر الكلمة في مقطعها الأول. وأخرى تنبر المقطع الأخير منها...كما يبرز في هذه الخرائط الدرس الواسع للمفردات. من حيث البنية والمترادفات المختلفة للمعنى الواحد واختلاف الألفاظ باختلاف المناطق اللغوية. ومقدار انتشار الكلمات في الأقطار والأقاليم. وغير ذلك؛ مما يتيح لنا معرفة الواقع اللغوي للغة من اللغات. سواء أكانت لغات فصحى أم مشتركة أم خاصة. أم لهجات اجتماعية. أم إقليمية. أم عاميات خاصة)(عبد التواب. رمضان. 1997 : 147).

لقد تجسد هذا المطمح العلمي في صناعة أطلس لساني جزائري في صورة تمظهرات علمية متعددة برزت بالأخص في المجال الأكاديمي؛ من خلال اشتغال عديد الدارسين. والفرق البحثية عبر مختلف ربوع الجزائر. في وضع خرائط لسانية متنوعة؛ تتبع الظواهر اللهجية الجزائرية في مجالات محددة؛ لا سيما تلك التي يمكن وصفها في هذا المجال بـ"الجزر اللغوية". الذي يؤشر على المناطق التي تبرز فيها الظواهر اللغوية. بشكل حاد ومتميز عن سواها في مناطق الساحل أو الهضاب أو الصحراء ومن أدنى المناطق الشرقية إلى أقصاها في الناحية الغربية. هذا لا يعني البتة أن عملية المسح الجغرافي اللغوي قد تحققت بالصورة المثلى؛ إذ لم نشهد إلى اليوم ظهور عمل علمي متكامل يرسم الخريطة اللغوية الجزائرية ويحدد معالمها بوضوح تام. إلا أن ذلك لا ينفي المراحل المتقدمة التي قطعت في هذا المشروع الوطني.

وجب التأكيد هنا على أن فكرة إنشاء أطلس لساني جزائري راودت كثيرا من الدارسين منذ أمد بعيد؛ فأُنجزتْ بحوثا علمية كثيرة اختصت بدراسة اللغة المحكية الجزائرية. واتصفت بالتنوع في مضامينها وطرائق انجازها ومجالها الزماني والمكاني؛ في عدد من الجامعات الوطنية. وإذا كانت بعض هذه البحوث وجدت سبيلها إلى النشر فإن بعضها الآخر ظل حبيس الرفوف في المكتبات الجامعية. وهو ما يحول دون الاستفادة منها. فلو أمكن للمهتمين بالأطلس اللساني الجزائري جمع هذه البحوث للاستفادة من نتائجها؛ وهذا بوضع بنك لهذه البحوث والدراسات. وهذه الدعوة باختصار تعني استثمار التكنولوجيات الحديثة في انجاز هذا المشروع الحضاري وتوظيفها بحيث يحقق أهدافه العلمية المرجوة. إن الانتقال من الورقية إلى الرقمنة من شأنه أن يعرّف بهذا الإرث اللغوي اللامادي.

E-ISSN - 2773-2886 / ISSN - 2773-2894

ويحفظه أمام التحديات المعاصرة التي تواجه مجتمعنا العربي في هذه الألفية الجديدة على الصعيد العام؛ فمن شأن(تكنولوجيا المعلومات بما توفره من وسائل عديدة في المجال اللغوي أن تتيح فرصا عديدة لانتشال العربية من أزمتها الراهنة واستعادة مجدها القديم. حتى تمارس دورها المحوري المنوط بها...وعلينا إدراك أن الانفجار المعلوماتي سلاح ذو حدين؛ فإما أن يكون مصدرا لزاد معرفي متوافر ومتجدد. وإما أن يكون عبئا ننسحق أمامه) (نبيل علي. 2005 : 323).

إن الحديث عن الانتقال من النشر الورقي إلى النشر الرقمي للأطلس اللساني الجزائري؛ يعني في واقع الحال: وضع الاستعمال اللغوي الجزائري المكتوب والمنطوق من خلال خرائط تبيِّن التَّوزيع الجغرافيّ للأنماط اللغويَّة- فصيحها وعاميها- واستخداماتها؛ في شكل تطبيقات رقمية وبرامج الكترونية متنوعة؛ تتيح لاستخدامه والاطلاع عليه على نطاق واسع. ومن ثمة تحويله من كتل معرفية ورقية ثابتة غير متاحة للجميع. إلى كتل معرفية رقمية ذات جودة عالية. وحركية فاعلة في مجتمع المعرفة الإنساني. قد تكون هذه التطبيقات في شكل صور لخرائط لغوية رقمية. تتحكم فيها التكنولوجيات الحديثة؛ على غرار تلك الخرائط الجيو فيزيائية التي نشاهدها يوميا في المؤتمرات العلمية. وقاعات البحث وشاشات التلفاز. وقد تكون في صورة منصات رقمية ثابتة؛ كتلك التي خصصت للمعاجم اللغوية العالمية. يلج إليها المهتمون والقراء. فتعرض لهم السجل اللغوي الاجتماعي في أصواته. وألفاظه. وتراكيبه. أو من خلال برامج الكترونية مثل الفيديوهات القصيرة أو المطولة التي ترصد مختلف ظواهر الاستعمال اللغوي اليومي. وربما تكون في شكل صور رقمية تكشف واقع مسميات الاستعمال اللغوي في الحياة اليومية للفرد الجزائري.

E-ISSN - 2773-2886 / ISSN - 2773-2894

4 . دور المعالجة الآلية ومنهجيتها في صناعة الأطلس اللساني الجزائري الرقمي

إن ظهور الحاسب الآلي مع منتصف القرن الماضي قد أدى إلى حدوث تغيير جذري في مجال الدراسات اللغوية القائمة على النصوص اللغوية؛ ودفعت ثورة المعلومات إلى تغيير أساليب العمل القائمة على المواد اللغوية بصورة جذرية. ولم تعد هناك حاجة إلى استخدام البطاقات في جمع المواد اللغوية وتصنيفها؛ حيث أصبح علماء اللغة ومصممو المعاجم يلجئون إلى استخدام الحاسب في تخزين كميات هائلة من النصوص ويمكنهم أيضا البحث عن المفردات أو العبارات أو الجمل وإظهارها على شاشة الحاسب بسرعة دون أدنى خطأ. والأكثر من ذلك أنه يمكن إعادة التصنيف والترتيب بناء على أكثر من أسلوب.. (خوانغ تشانغ نينغ. 2016 : 49). انطلاقا من ذلك كله فإنه يؤمل من المعالجة الآلية الحاسوبية للأطلس اللساني الجزائري تحقيق عديد الانجازات المعرفية النظرية والتطبيقية؛ والتي يمكن أن نوجز بعضها فيما يأتي:

1. سعة التخزين؛ فالمعالجة الآلية للمعطيات اللغوية لهذا الأطلس في الحاسب الآلي تمكننا من حوسبة عدد غير متناهي من المواد اللغوية التي تم جمعها لهذا الأطلس؛ وهذا ما أوضحه مازن الوعر: حينما قارن بين الذاكرة الإنسانية ذات الصفات المحدودة والقصيرة. فهي تعمل على أساس من النظام القصير؛ فيما تعمل ذاكرة الحاسب الآلي المركبة على أساس من النظام الثابت والطويل الأمد (الوعر. مازن. 1983: 13).

2. سرعة العمل العلمي؛ ففي المعالجة الآلية للأطلس اللساني نكسب رهان الوقت تنظيرا ودراسة ومطالعة؛ لأن السيرورة الزمنية متحكم فيها بحيث يتم إنجاز أي عمل علمي في حيز وجيز.

3. الموضوعية والدقة في الانجاز؛ هي من الجوانب الأساسة التي تميز المعالجة الآلية الموضوعية؛ فالعمل ينجز بطريقة سلسلة لا مجال فيها للذاتية في الاختيار والانجاز. ومن جانب آخر تعرض للظواهر اللغوية بالدقة المتناهية؛ انطلاقا من المعطيات المتاحة للحاسوب.

4. تعدد وظائف العمل العلمي؛ ففي سياق الاشتغال بالأطلس اللساني تمكن المعالجة الآلية من تنويع وظائف العمل؛ فللدارس انجاز نشاطين متكاملين أو أكثر في الآن نفسه. من نحو البناء

والتركيب والتحليل والفهم في الآن نفسه؛ وهو ما لم يكن متاحا مع العمل البشري اليدوي (الوعر. مازن. 1983: 13).

ويفرق الدارسون بين نوعين رئيسيين من المنهجيات المتبعة في مهام المعالجة اللغوية؛ هاتان المنهجيتان يمكن توضيحهما فيما يأتي(ديانا ما يندارد. 2019: 41):

1. أحدهما منهجية قائمة على المعرفة أو القائمة على القواعد؛ تعد من الأساليب التقليديةً بصفة عامة. وتستند هذه المنهجية على قواعد مكتوبة يدويا. وتجري كتابة هذه القواعد عادة على يد متخصصين في مجال معالجة اللغات الطبيعية. وتتطلب معرفة قواعد اللغة والمهارات اللغوية. فضلا عن امتلاك ملكة البديهة. تكون هذه المنهجيات ذات فائدة أكبر إن أمكن تعريف المهمة بسهولة بواسطة القواعد المنهجية القائمة على المعرفة (على سبيل المثال قاعدة: «الاسم الصحيح -في اللغة الإنجليزية- يبدأ دائًما بحرف كبير.)» وفي العادة. يمكن استثناء هذه القواعد بسهولة.

عندما لا تنطبق القاعدة اللغوية بشكل مباشر تولد هذه المنهجية إشكالية أكبر من السابق (على سبيل المثال: في تغريدات تويتر غالبا لا يستخدم الناس الأحرف الكبيرة لكتابة الأسماء الصحيحة -في اللغة الإنجليزية-.) من بين المزايا الكبيرة للمنهجية القائمة على المعرفة السهولة الكبيرة في فهم النتائج. عندما يتعرف النظام على شيء ما بشكل غير صحيح. يكون بوسع المطور التحقق من القواعد وتحديد سبب حدوث الخطأ. ومن ثم يحتمل أن يكون بمقدوره تصحيح القواعد أو كتابة قواعد إضافية لحل المشكلة. ومع ذلك. يمكن أن تستهلك عملية كتابة القواعد الكثير من الوقت. وفي حال حدوث تغيير في المهمة. فقدّ يضطر المطور إلى إعادة كتابة العديد من القواعد.

2. منهجية مبنية على التعلم. أو منهجيات تعلم الآلة تحظى بشعبية أكبر في الآونة الأخيرة مع ظهور أجهزة قويةً ومتطورة. وأيضا بسبب عدم وجود ضرورة لامتلاك خبرة في المجال المعنِّي أو امتلاك معرفة لغوية. ولذلك أصبح بالإمكان أن ننشئ نظاما خاضعا للإشراف بسرعة كبيرة إذا توفرت بيانات تدريبية كافية. وبوسعنا الحصول على نتائج معقولة بعد تدريب محدود جدا. غير أن الحصول على بيانات تدريبية كافية أو إنشاءها غالبا ما يطرح إشكالية كبيرةً للغاية ويستغرق وقتّا طويلا. ولا سيما إذا كان لا بد من القيام بهذه العملية يدويا.

يعني هذا الاعتماد على بيانات التدريب أيضا أن التكيف مع أنواع جديدة من النصوص أو المجالات أو اللغات سيكون مكلفا على الأرجح؛ حيث يتطلب توفر قدر كبير من بيانات التدريب الجديدة.. لذا فإن القواعد التي يكون البشر قادرين على قراءتها عادة ما تكون أسهل في التكيف مع اللغات وأنواع النصوص الجديدة مقارنة بتلك المبنية على أساس النماذج الإحصائية. كما يمكن معالجة مشكلة توفر بيانات التدريب الكافية عبر الدمج بين التعلم الآلي. والطرق غير الخاضعة أو شبه الخاضعة للإشراف. مع العلم أنها عادة ما تعطي نتائج أقل دقة مقارنة بنتائج التعلم الخاضع للإشراف. وعند المقارنة بين المنهجيتين نستخلص ما هو موضح في الجدول الأتي(ديانا ما ينارد. 2019: 41):

أنظمة التعلم الآلي	المنهج القائم على المعرفة
تستخدم علم الإحصاء أو أساليب التعلم الآلي الأخرى	تقوم على قواعد مكتوبة يدويا
لا يتعين على المطورين أن يكونوا على دراية بمعالجة اللغات الطبيعية	جرى تطويرها على يد متخصصين بمعالجة اللغات الطبيعية
تتطلب كميات ضخمة من البيانات التدريبية	تستغل ملكة البديهة البشرية
يصعب فهم أسباب وقوع الأخطاء	نتائج سهلة الاستيعاب
عملية التطوير سهلة وسريعة	قد تستهلك عملية التطوير وقتا طويلا للغاية
قد تتطلب التغييرات عملية إعادة إضافة تعليقات وشروحات	قد تتطلب التغييرات إعادة كتابة القواعد

الجدول 01 : ملخص المنهج القائم على المعرفة في مقابل المنهج القائم على التعلم الآلي في معالجة اللغات الطبيعية

وقد سعت المعالجة الآلية للغة العربية إلى بناء قواعد بيانات بديلة عن الورقية في التدوين والكتابة والتأليف؛ وهو ما يمكن استثماره في بناء قاعدة للأطلس اللساني الجزائري الرقمي؛(حيث سرعة الاسترجاع. وسهولة الإدارة. والتحديث والتطوير والتغيير. وسرية البيانات وأمنها. وقد انتهت المحاولات الدائبة لتطوير هذا الحقل البحثي إلى حلول قواعد البيانات (المعلومات/المعطيات) محل الملفات التقليدية؛ وبذلك صار ممكنا تجنب نواقص الملفات التقليدية في معالجة البيانات وحفظها واسترجاعها. وصار ممكنا الحصول على بيانات متعددة بخطوات أقل وبسهولة أكثر من الملفات التقليدية) (العناتي. وليد. 2012: 291).

5 . مداخل استثمار المعالجة الآلية في صناعة أطلس لساني جزائري رقمي

تتعدد أغراض المعالجة الآلية للغة؛ بوصفها أداة ذكية قادرة على خدمة الدرس اللغوي من خلال ما تتيحها ببرمجياتها المتنوعة التي تتناغم ومستويات النظام اللغوي؛ والتي يمكن استثمارها في صناعة أطلس لساني جزائري رقمي؛ ومعلوم أن المعالجة الآلية للّغة العربية بشكلٍ عامٍ تهدف إلى:

1 . دمج المعرفة اللّسانية بالحاسوب.

2 . الإبداع في طرق التفكير اللّساني الحاسوبي.

3 . تحقيق التكامل بين مختلف العلوم اللّغوية والعلوم الرياضية والعلوم الحاسوبية.

4 . إنتاج المعرفة الكلّية لتوحيد الرؤى الجديدة.

5 . الاطلاع على التّجارب المحلّية والإقليمية والعالمية الخاصة بهذا المجال. وبالاعتماد على الأبحاث البينية التي حاولت التأسيس لعلاقة وطيدة بين المفاهيم اللسانية والحاسوبية وآلياتها الرقمية وتسخيرها لخدمة اللغة العربية. برزت مجالات معرفية لسانية متعددة لهذه العلاقة كان من أهمها على الإطلاق حوسبة المعجم العربي. الذي ظهرت له برامج رقمية متنوعة. بل ووضعت له منصات الكترونية عالمية تمكن الدارس من الولوج لها في كل الأوقات. في هذا الإطار يمكننا إيراد ما ذهب إليه علماء اللسانيات الحاسوبية من أوجه استخدام المعالجة الآلية؛ وهي كالأتي:

E-ISSN - 2773-2886 / ISSN - 2773-2894

أ. استخدام المعالجة الآلية في الإحصاء اللغوي

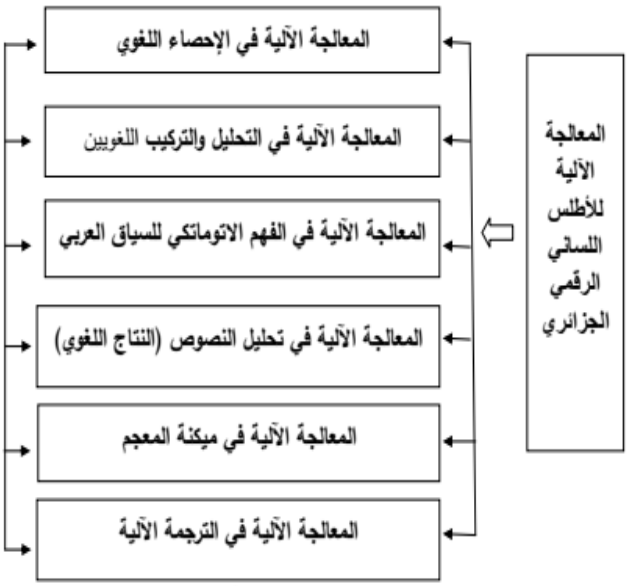
ب. استخدام المعالجة الآلية في التحليل والتركيب اللغويين

ج. استخدام المعالجة الآلية في الفهم الأتوماتيكي للسياق العربي

د. استخدام المعالجة الآلية في تحليل النصوص (النتاج اللغوي)

هـ. استخدام المعالجة الآلية في ميكنة المعجم

و. استخدام المعالجة الآلية في الترجمة الآلية(بن عربية. راضية.2017: 251).



خطاطة: المداخل الحاسوبية اللغوية للأطلس اللساني الرقمي

تحمل هذه الخطاطة ببساطة جوانب الاشتغال على الأطلس اللساني بشكل حاسوبي؛ وتبرز العديد من العمليات؛ التي يمكن أن ينهض الحاسوب بها في خدمة هذه الأطالس. وتوظيف معالجتها الآلية التي تقوم بدمج المعارف اللغوية في الحاسوب في صورة آلية. وبرمجتها وفق لغة اصطناعية؛ (ولغة البرمجة بكل بساطة هي اللغة الصورية المترجمة لواقع اللغة الطبيعية؛ بحيث يمكن للآلة فهمها

وقراءاتها والتعامل معها؛ وتبنى أساسا على وحدة الخوارزم وحدة أساسية لكل تعامل حاسوبي مع معطيات اللغات الطبيعية) (نبيل علي. 2005 : 131).

وإذا نظرنا إلى المدخل الأول لأوجه المعالجة الآلية للأطلس اللساني الجزائري - الإحصاء اللغوي - نجد الاستفادة من هذه المعالجة الآلية في صناعة هذا الأطلس عظيمة؛ ذلك أن للغة العربية ذات منطق إحصائي؛ فاللغة العربية لغة طبيعية تقوم في بنائها التركيبي على مجموعة من المعادلات الرياضية الصرفة؛ قابلة للإحصاء. وقد بينت عمليات التشريح الإكلينيكي للغة العربية تميزها سمة القابلية للإحصاء في بياناتها المختلفة؛ الصوتية. والصرفية. والتركيبية. إن المعالجة الآلية أداة فاعلة للإحصاء اللغوي لمختلف قضايا الأطلس اللساني؛ ذلك أن تعدد الظواهر اللغوية وتلازمها؛ يتطلب جهدا كبيرا من الباحثين يمكن للرقمنة أن تختزله تبويبا وتصنيفا وترتيبا. ولعل من أهم المجالات العملية في استعمال الحاسب والمعالجة الآلية في رقمنة الأطلس اللساني الجزائري ما يأتي: (الوعر. مازن. 1983: 13):

(1). معالجة النصوص أو تنسيق الكلمات word processing؛ ويقدم لها هذا النوع من البرمجيات إمكانات الكتابة وتوابعها مثل التدقيق الإملائي. والتدقيق النحوي. والمساهمة في دقة التعبير المعجمي من خلال معجم المترادفات أو المكنز thesaurus. إضافة إلى إمكانية إحصاء عدد ألفاظ النص وترتيبها sort (عند معاملتها على شكل فقرات). والبحث عنها في النص كذلك find (كما نجد في برنامج Word). وغير ذلك من الإمكانات. مما يوفر على المستخدم الكثير من الجهد والوقت.

(2). عرض النصوص لأغراض الشرح والمناقشة الجماعية والتدريب (برنامج power point مثلا).

(3). استخدام الشابكة (الإنترنت) لأغراض مختلفة. مثل البحث عن المعلومات. والحصول على نسخ من بعض المطبوعات. والبحث المعجمي. والوصول إلى المدونات المتاحة على الشابكة.

(4). برامج قراءة النصوص المكتوبة: حيث نستفيد من تقنيات عدة منها:

E-ISSN - 2773-2886 / ISSN - 2773-2894

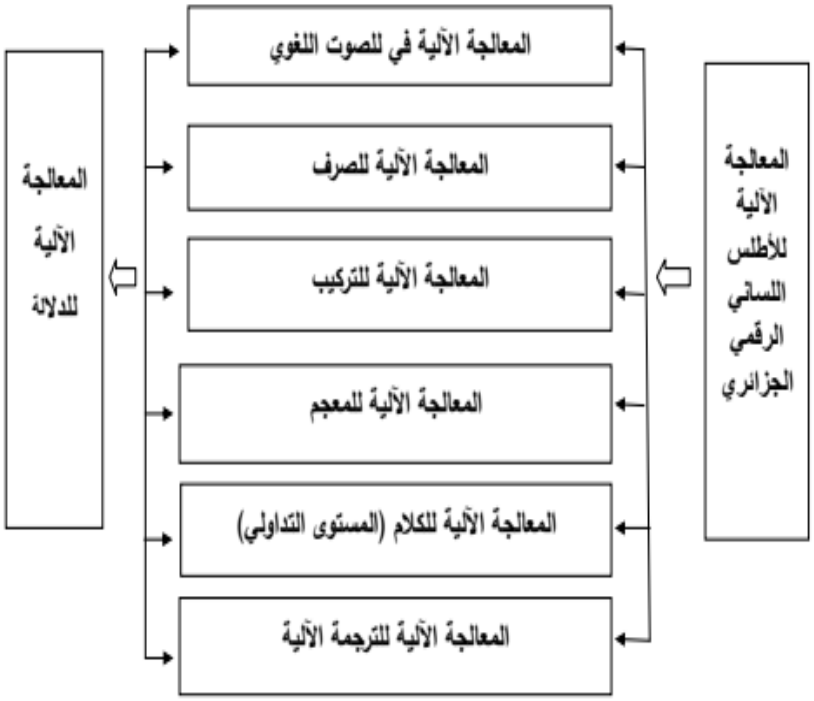
(أ). التعرف على المحارف البصرية optical character recognition OCR التي تحول النصوص المكتوبة إلى مادة رقمية, ومن أهم أوجه الاستفادة من هذا النوع من البرمجيات؛ تخزين النصوص المطبوعة في قواعد المعلومات والمدونات اللغوية .

(ب) برنامج إنتاج الكلام أو توليف الكلام speech synthesis لتحويل النص الرقمي إلى كلام منطوق.

(5). التعرف على النصوص الشفوية: ويعتمد على تقنية التعرف على الكلام speech recognition. ثم(أ) تحويله إلى نص مكتوب في حالة الإملاء الآلي أو(ب) تنفيذ الأوامر المنطوقة. أو(ج) "فهمه" في حالات الاستعلام الشفوي مثلا. ويستخدم الأخير تقنية أخرى أيضا هي تقنية الذكاء الاصطناعي(artificial intelligence (AI.

وتبرز الأهمية البالغة للمعالجة الآلية للأطلس اللساني الرقمي في جوانب أخرى من نحو؛ استنطاق النصوص. وفي إدخال النصوص وتخزينها بالمسح الضوئي. وفي الّتحليل والتركيب الّنحوي والّصرفي. والإعراب الآلي. وفي المعاجم الالكترونية. وفي الترجمة الآلية. وفي بنوك المصطلحات والمعلومات. وفي الفهرسة. وفي قواعد المعارف. وفي التحليل الإحصائي....(محمود إسماعيل صالح. 2019).

ليس ذلك فحسب بل إن هذا الدور يتعاظم مع دمج مستويات النظام اللغوي ضمن عمل الحاسب الآلي على مستوى الصوت. والصرف. والتركيب. والمعجم. وفي المستوى الكلامي التداولي؛ وبشكل أوسع في الترجمة الآلية لمختلف معارف هذا الأطلس اللساني؛ هذه المستويات التي يبدو تفاعلها مع بعضها في إطار تكاملي لا ينفصل أحدهما عن الآخر وهو ما يمكن التمثيل له بهذه الخطاطة البسيطة:



خطاطة: المعالجة الآلية للنظام اللغوي للأطلس اللساني الرقمي الجزائري

6 . خاتمة

نخلص في ختام هذه المداخلة إلى أن المعالجة الآلية للأطلس اللساني الجزائري مطمح علمي ومنجز لساني حضاري مميز في مطلع هذه الألفية الجديدة؛ يحقق الإضافة لدراسة واقعنا اللغوي الجزائري إذا تحقق بالصورة المثلى التي ننشدها؛ فيما تعكف على تقديمه فرق البحث اللسانية المتخصصة في مختلف المخابر ومراكز البحوث الوطنية؛ لتهدف من خلاله إلى رقمنة التراث اللغوي الجزائري في البرامج المعلوماتية المعاصرة. وإتاحة الاطلاع عليها للباحثين والمهتمين على نطاق واسع من خلال الشابكة (الانترنت)؛ في صورة أرضيات حاسوبية. وبرامج رقمية متنوعة؛ ومن أهم ما توصلت إليه هذه المداخلة:

- الرقمنة والمعلوماتية. وحوسبة مختلف أنساق المعرفة من أوضح سمات هذا العصر؛ لما يمتلكه الحاسوب من إمكانات هائلة في استقبال البيانات. والمعلومات ومعالجتها وتخزينها. وتقديمها بسرعة ودقة.

- والدعوة إلى تأسيس أطلس لساني رقميٍّ؛ في واقع الحال هو وضع الاستعمال اللغوي الجزائري المكتوب والمنطوق في صورة خرائط تبيّن التَّوزيع الجغرافيّ للأنماط اللغويَّة واستخداماتها؛ في شكل تطبيقات رقمية وبرامج الكترونية متنوعة؛ تتيح استخدامه والاطلاع عليه على نطاق واسع.

- يمكن تصور مداخل حاسوبية لغوية متعددة للمعالجة الآلية للأطلس اللساني الجزائري الرقمي؛ في الإحصاء اللغوي. وفي التحليل والتركيب اللغويين. وفي الفهم الأتوماتيكي للسياق العربي. وفي تحليل النصوص (النتاج اللغوي). وفي ميكنة المعجم. وفي الترجمة الآلية.

- تبرز الأهمية البالغة للمعالجة الآلية للأطلس اللساني الرقمي في جوانب متعددة من نحو؛ استنطاق النصوص. وفي إدخال النصوص وتخزينها بالمسح الضوئي. وفي الّتحليل الّنحوي. والّصرفي. والإعراب الآلي. وفي المعاجم الالكترونية. وفي بنوك المصطلحات والمعلومات. وفي الفهرسة. وفي قواعد المعارف.

ولعل من أهم التوصيات التي نوجه أنظار الدارسين إليها؛ بناء أرضيات رقمية كبرى لمشروع الأطلس اللساني الجزائري. ووضع قاعدة بيانات لغوية جغرافية رقمية للمنطوق الجزائري. وصناعة معاجم اللغات واللهجات الجزائرية. وإعداد دليل وطني للباحثين في اللهجات الجزائرية.

7 . قائمة المصادر والمراجع

- بن شاشة. فارس.(2008).المعالجة الآلية للغة العربية؛ إنشاء نموذج لساني صرفي للفعل العربي. مذكرة ماجستير. كلية العلوم الإنسانية والاجتماعية. جامعة الجزائر. الجزائر.

- بن عربية. راضية. (2017) محاضرات في اللسانيات الحاسوبية. ط01. ألفا للوثائق والنشر. الجزائر.

- بن عربية. راضية.(2017). حوسبة النظام اللغوي العربي.- المعجم الآلي عند البروفسور عبد الرحمان الحاج صالح أنموذجا ضمن كتاب ازدهار اللغة العربية بين الماضي والحاضر. ط01. المجلس الأعلى للغة العربية. الجزائر.

- براهمي. إبراهيم.(2018). نَحْو أَطْلَسٍ لُغَوِيٍ جَزَائِرِيٍ جَامِعٍ- قِرَاءَةٌ فِي المَفَاهِيمِ والآلِياتِ والأَهَدَافِ-. الملتقى الدولي: صناعة الأطالس اللسانية -الآليات والتجارب. وحدة البحث اللساني وقضايا اللغة العربية في الجزائر. ورقلة- يومي: 11 و12 ديسمبر 2018.

- ديانا ما نارد. كالينا بونتشيفا. ايزابيل أوغنشتاين. (2019). معالجة اللغات الطبيعية دلاليا. تر: خالد الميمان. ط01. مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية. الرياض. السعودية.

- خوانغ تشانغ نينغ. (2016). علم الذخائر اللغوية. تر: هشام موسى المالكي. ط01. المركز القومي للترجمة. القاهرة. مصر.

- السالمي. علاء عبد الرازق.(2002). تكنولوجيا المعلومات. ط02. دار المناهج للنشر والتوزيع. عمان. الأردن.

- مختار عبد الخالق عبد الله.(2007). تعليم اللغة العربية باستخدام الحاسوب. ط01. دار العلم والإيمان. الإسكندرية. مصر.

- نبيل علي. نادية حجازي.(2005). الفجوة الرقمية. رؤية عربية لمجتمع المعرفة. عالم المعرفة. العدد:318. الكويت.

- مجموعة من المؤلفين.(2014). اللسانيات وإعادة البناء؛ وقائع الندوة العلمية الدولية الثالثة للسانيات. المنعقدة بتاريخ: 10. 11. 12 أفريل 2014م. كلية الآداب والفنون والإنسانيات. جامعة منوبة. ط01. تونس.

- العناتي. وليد.(2012) العربية في اللسانيات التطبيقية. ط01. دار كنوز المعرفة. عمان. الأردن.

 E-ISSN - 2773-2886 / ISSN - 2773-2894

- غسان مراد.(2019). المعالجة الآلية للغة العربية: الإمكانات والفوائد والآفاق. مجلة الحياة. الرابط تاريخ الاطلاع : 02 أوت 2019م من خلال الرابط الالكتروني:

www.alhayat.com/article/1241085

- عمر. أحمد مختار. (2008). معجم اللغة العربية المعاصرة. ط01. عالم الكتب. القاهرة. مصر.

- عبد التواب. رمضان.(1997). المدخل إلى علم اللغة. ط03. مكتبة الخانجي. القاهرة. مصر.

- الواعر. مازن. (1983). اللسانيات والعلم والتكنولوجيا؛ نحو تعريب موحد للسانيات التطبيقية وبرمجتها في الحاسبات الالكترونية. مجلة اللسان العربي. العدد22. مكتب تنسيق التعريب. الرباط. المغرب.

- نبيل علي. (1988). اللغة العربية والحاسوب. ط01. مؤسسة تعريب. القاهرة. مصر.

- محمود إسماعيل صالح. (2019). الحاسوب واللغة والبحث اللغوي. تاريخ الاطلاع عليه: 18أوت2019م من خلال الرابط الالكتروني:

http://dr-mahmoud-ismail-saleh.blogspot.com/2013/12

- Georges Mounin, (2004), Dictionnaire de Linguistique, GUADRIGE/PUF, édition04, , paris, France.