

حوسبة اللغة العربية في ضوء المتجدد المعلوماتي "مشروع الذخيرة اللغوية الأستاذ عبد الرحمن حاج صالح "أنموذجا"

Digitization of the Arabic language in the light of IT innovations Corpus: ADHAKHIRA
Language Project Abderrahmane Hadj Salah

الدكتور: ابراهيمي بوداود

brahimtc@yahoo.fr

المركز الجامعي أحمد زبانة – غليزان (الجزائر)

تاريخ النشر: 2019/03/19

تاريخ القبول: 2019/02/13

تاريخ الإرسال: 2019/01/23

ملخص: تُعد المعالجة الآلية للغة مجالا معرفيا تجاوز حدود الذكاء الفطري للإنسان، لبراد فضاء الذكاء الاصطناعي، حيث تصل الكفاءة العلمية إلى ذروة سنامها، وبذلك تهيأ للغة أفقا إجرائيا ينأى بها عن المعالجة المجردة التي تسترشد بالكفاءة العلمية للدارس لتمتاع من الوضع التقني المتجدد الذي أفرزه الزمن التكنولوجي، الذي استجابات النظرية اللسانية لإغراءاته، في مسعاها لدخول عوالم التقنية ومراودة فضاء الذكاء الاصطناعي، المعتمد في اللغات الطبيعية والاستفادة من الإمكانيات التي يتيحها الحاسب الآلي، وتجاوز تلك النظرة التقليدية التي تحصر أسباب استخدامه في سرعته الفائقة.

ومن هنا، تأتي إشكالية البحث لتسلط الضوء على أنموذج إجرائي رائد في حقل حوسبة اللغة العربية، رام صاحبه-المرحوم الدكتور عبد الرحمن حاج صالح- إنشاء مدونة رقمية انطولوجية أو ما اصطلح عليه بمشروع الذخيرة اللغوية، وهو إنجاز ينهض على جمع الأنساق اللغوية في أسبقها الاستعمالية. ومن ثمة، تشتغل البرامج الحوسبية (المحلات) المُعدة على تحليل المادة اللغوية تحليلًا صوتيًا ومعجميًا ودلاليًا. وننتهي في الأخير إلى تقديم أنموذج تطبيقي لبرنامج WordNet نشتغل ضمنه على إيجاد المقابلات الدلالية للصيغ اللفظية في الجزائر. الكلمات المفتاحية: الحوسبة؛ اللسانيات؛ الأنطولوجيا؛ المحلات؛ الرقمنة؛ الثنائية الرقمية؛ الكودة؛ الدلالة؛ السياق؛ الجمع؛ الخاصية؛ الجرافيم؛ الفونيم؛ التمثيل.

Abstract : The automatic processing of language as a field of knowledge has considerably outstripped the limits of innate intelligence in man to brew the sphere of artificial intelligence, or scientific competence unquestionably reaches its climax, thus, it enters a way to an operating horizon for the language by penetrating into the universe of the technicality, already used in the natural languages and to benefit from the possibilities which the computer offers, and the surpass the archaic vision which limits the reasons from its use only processing speed. Indeed, our problematic attempts to shed lights on the current paradigm in the field o computer-assisted processing, whose forerunner, the term of the language arsenal.

While ending, finally, in the presentation, an applied model of Word Net software that we will try t highlight semantics dialects equivalences in Algeria.

Key words : Informatory, linguistics, anthology, numeration, binary, semantics, context, character, grapheme, code.

تصدير

انقادت إجرائية الدرس اللساني الحديث نحو مصير التمثل القرائي للبرنامج اللغوي البنوي، الذي انفتح على آفاق نوعية وضعت النظام اللغوي في موقع متميز من المقاربات الرقمية، حيث انبرت لمواجهة التحديات التي أفرزتها عقلية المرحلة التقليدية، التي أوغلت في الإنصات لنبض البنى التركيبية الجاهزة وسلطة القوالب النحوية الجامدة، فكان أن انعتقت الدراسة اللغوية من بوتقة التحليل النمطي لترتاد آفاق تحليلية اخترقت حدود العلوم المادية لتصبح الحواجز بينها وبين علم الحاسوب حواجز وهمية مائعة، ساهمت في إدخال الدرس اللغوي في سياق فكر المرحلة الجديدة.

إزاء هذا الوضع تعزز الطرح اللساني بآليات وتقنيات إجرائية رقمية عضدت من مكنة تثبيط الكثير من الصعوبات التي كانت تفرضها طبيعة البحث في اللغة، بوصفها ظاهرة ترتين إلى التجريد الحسي وتؤول

إلى الاستغراق اللامتناهي الذي يؤديه الفكر، إذ لم يعد في وسعها التملص من سلطة المسلمة الرياضية والاستلزمات البرمجية، التي تحدد النتيجة، فغدا البحث في حقيقة التشكلات الصورية لتراتبية الوحدات اللسانية وصنافات تعاقباتها الخطية، مقابل التمثلات الذهنية لوظائفها المدلولاتية وتبدلاتها الاستعملية مطلباً جوهرياً يستمد شرعيته من تلك المقاربة للعمليات المنطقية اللغوية التي تؤدها المنظومة اللسانية في العقل البشري، ضمن نسيج من العلائق المتداخلة تستدعي بدورها التقصي والتمحيص وفقاً لأبجديات تحليلية تستشرف عتبات الاستقصاء الرياضي والمعلوماتي.

فلاشك أن الوقائع التجسيدية للمدرجات اللسانية تتمظهر في هيئة حسية تتشكل ملامحها في صورة أجزاء مترتبة من المنطوق البشري، مادتها الصوت اللغوي. ومن ثمة فإن أية محاولة تدعو إلى تشفير نسيج اللغة لا تعنص بالمنطق الذي يحكم المستوى الرمزي لا تعدو أن تكون إلا ضرباً من العبث المتهجي بآليات البرنامج اللغوي. وعلى هذا الأساس "أطبق جمهور العلماء على أن اللسانيات عامة، هي أكثر العلوم الإنسانية دخولا إلى فردوس العلوم المنضبطة"¹. وذلك من منطلق مطاولتها، لتلك المناهج الاستنباطية العقلية التي ترتبهن إلى سلميات المنطق الذي يلفظ التناقض بين مقدماته ونتائجه.

وبالتساند إلى هذا الاستقطاب النظري، نلمح تفريعا نوعيا في سياق الأطروحات اللسانية انبثق من غمرة المد المنهج الرياضي، بوصفه أنموذجا مؤطرا يمتلك الجهاز الإجرائي لإدغام المستويات اللغوية وتوجيه مسارها نحو دروب التثبت الرقمي اليقيني، كونه يركن إلى تخوم معرفية من شأنها أن تترفع بمستويات الطرح فيه إلى درجات تتأبى الانصياع لمنطق الذاتية والريب، بدءاً من مرحلة تشكله الأولى وصولاً إلى مرحلة جمعه وإدراكه ثم استيعابه، وكلها مراحل تستدعي تكاملاً معرفياً يرتد إلى العلاقة الصميمية التي تربط بين البرمجة الرقمية لمعطيات للنسق اللغوي -الوحدة اللغوية- والوظيفة المنطقية لها.

اللسانيات الحاسوبية العربية بين التصور النظري والآلية الإجرائية:

ترتبهن أبجدية التحليل اللساني إلى برنامج لغوي شمولي، يركن إلى ملابسات الفعل التواصلي ضمن منظومة اجتماعية، تخضع فيها اللغة لمنطق تراتبي ينهض على سلمية بنائية يتصدرها المعلم الصوتي بوصفه عتبة جوهريّة، تمخض عن تفاعل جزئياتها نسيج من العلائق التركيبية التي انبثقت من بؤرة الصوت فالمفردة فالتركيب ضمن مدار تجاذبي يشغل على استقطاب حيثيات البناء اللغوي وفقاً لشرائط دلالية تسيّر مشهيدة الأنساق اللسانية.

ولما كانت الإستراتيجية البنيوية للنظام اللغوي، تتكئ على معطيات السند العيني المادي لأنساق اللغة المغلقة، كونها خصيصة ملزمة للتحليل الآني السنكروني التي تستدعي الانطلاق منها لبناء معالم الخطاب اللساني، فقد شغلت الدراسات الحاسوبية والرقمية مكاناً رائداً من حيز المقاربات المعاصرة للغة، معلنة عن انبثاق توجه تحليلي مغاير لفكر المرحلة السابقة، وذلك عبر تجاوز خطية التمثل التعاقبي لأطروحات الدرس التراثي واقتحام الحدود الإجرائية التي تمخضت عن الحوار الآني بين علم اللسان والعلوم المادية المحوسبة، حيث امتحت الحواجز الفاصلة بينهما في ظل الإغراءات العلمية التي أفرزها الزمن التكنولوجي.

وفي غمرة هذا المد الحداثي تغيرت نبرة الخطاب اللساني، وتجردت من بقايا الحس الذوقي الذي استغرق في لامتناهيات التدافع الفكري والنظري ردحا من الزمن في غياب المرجعيات التقنية، مما استوجب الاحتكام إلى أبجديات الثقافة الانطباعية والتعويل على الأثر التعليلي للمسائل اللغوية التي أدت إلى تكريس جملة من المسالك المعيارية التي انغلقت على قناعات الماضي ومحدداته الاستمولوجية، لتنتفح المؤسسة اللغوية العربية -لاحقا- على دروب مغايرة جراء الانعطاف الإجرائي الطارئ على المنظومة اللغوية بفعل الزخم المعرفي الذي انفتحت عليه الدراسات اللسانية الحديثة، فكان أن تهيأت لاحتواء مشروع البعث العلمي الذي دعا إليه التوجه البنوي بالانعتاق من أسر التوجه التسيني الذاتي والاسترشاد بمكتسبات الزمن التكنولوجي. إزاء هذا التحول النوعي في إستراتيجية البحث اللساني، انبثقت الأسئلة عن سبل الإحاطة بتفاصيل الكينونة اللسانية بمختلف مظهراتها الترميزية وتجلياتها الدلالية، فكان أن استشرفت الدراسة اللسانية عتبات البحث الحاسوبي لتقدم مقاربات علمية تنأى عن نمطية التوجه التراثي ذي الطابع العمومي الشمولي، إذ تبلورت التصورات الأولية وفقا لآلية إجرائية تعنى بضبط الظاهرة اللغوية بمعزل عن المرجعيات الخلفية، والتصورات القبلية مشيرة إلى إمكانية تحول علم اللغة الحديث من موقف الند للدرس القديم باعتبار السيرة الزمنية إلى موقف النصير، مما يعين على التثبت مما أفرزته العقلية اللغوية القديمة وإحلال أطروحاتها في موقع متميز من الفضاء اللساني العربي بدلا من دحض تصوراتها في إطار القطيعة الاستمولوجية بين التراث والحداثة وهدم مسلماتها.

ومن هنا، اشتغلت الدراسة على مد جسور التواصل بين الدرس اللغوي التراثي بوصفها لغة طبيعية وما استقرت عليه المنظومة اللسانية الحاسوبية الحداثية بالتطلع إلى نسقية مفتوحة تتكى على عتبة التراحم المعرفي في إطار مشروع تقني يهدف إلى إخراج حقل علوم اللغة من بوتقة التنظير المسلم إلى رحاب يرتاد آفاق المعاينة الآلية وفقا لازدواجية منهجية تتخذ من الغرافام graphème مقابل الشفرة code بؤرة تحليلية تمتاح من خصوصية المدونة اللسانية العربية بوصفها سندا مرجعيا، قصد "التأسيس لقاعدة بيانات لتنظيم الثروة اللغوية في سائر مستوياتها: الصوت والدلالة والصرف والنحو في صورتها المنطوقة والمكتوبة². ومن ثمة، الانفتاح على إفرازات الأنظمة والبرمجيات باعتبارها سننا آلية تعين على التثبت اليقيني من نتائج التحليل والتععيد.

وانطلاقا من أن المشروع الذي يعتمد التحليل الحاسوبي استراتيجي إجرائي له، يستوجب حتما الاتكاء على قاعدة بيانات ذات معالم لسانية وحاسوبية ورياضية وهندسية، ولذلك اقترح الدكتور عبد الرحمن الحاج صالح، تأسيس دراسة تهض على الحوار المعرفي بين مختلف الميادين³، حتى يتهيأ للجميع اكتساب "لغة فنية مشتركة لطول الاجتماع والمحاورة⁴، حيث استقطب مشروع "العلاج الآلي للغة" عددا من الرياضيين والمهندسين في المعلومات الذين تم تزويدهم بدروس مكثفة في اللسانيات واللسانيات العربية، في حين تدرب حاملي شهادة الليسانس في الأدب العربي على تقنيات البرمجة الآلية ودروسا في الرياضيات التطبيقية والحاسوبيات⁵. وبذلك تمكّن الدكتور عبد الرحمن الحاج صالح من الانفلات من مأزق طغيان "الجانب الحاسوبي على الجانب اللساني، أو العكس⁶.

طرح عبد الرحمن حاج صالح في التأسيس لمبادئ اللسانية الحاسوبية العربية

انبثقت مشروعية التنافذ المعرفي بين الدرس اللساني والحقول المعرفية التطبيقية، إثر التحول المنهجي الذي شهدته الدراسات اللسانية الحديثة، حيث انعطفت بمساره عن المسلك التقليدي لتراود فضاء تحليليا يستشرف من خلاله الدارس أفقا علميا، ينأى بالدراسة عن سيطرة الذوق اللغوي، والاستسلام لمحدداته الذاتية، فقد " دأبت الدراسات اللسانية منذ أن تشكلت مرتكزاتها النظرية والإجرائية في بداية القرن العشرين على ضبط الإواليات التنظيمية للأنساق اللغوية المختلفة [...]، حيث تركزت اهتماماتها على وصف بنيات اللغات، وتقنين أنماط اشتغال مكوناتها [...] وفق استراتيجيات بحث جديدة تؤسس لنفسها موقعا علميا متميزا تعيد من خلاله رسم حدود التقاطع القائم بينها وبين المعلومات، والمنطق والسيكولوجيا والذكاء الاصطناعي داخل المجال الأرحب للعلوم المعرفية"⁷.

وبذلك ألفينا انبثاق معرفة لسانية يزدوج فيها الذكاء الفطري بالذكاء الاصطناعي، الذي يعد "إحدى أهم الخطوات العلمية للانتقال من اللغات الطبيعية إلى اللغات الاصطناعية باعتماد أنماط متعددة للتقييس. لقد مثل الذكاء الاصطناعي بالنسبة للعديد من الباحثين الجانب الهندسي للغة"⁸، فإذا كان مفهوم الهندسة يتحدد بوصفه فن السيطرة على النظم المعقدة، فإن مفهوم هندسة اللغة ينهض على السيطرة على النظام اللغوي عبر وسائط وآليات إجرائية تستقي ثوابتها من الإفرازات العلمية للسانيات الحاسوبية. وقد حاول الدكتور "عبد الرحمن الحاج صالح" -رحمه الله- اختراق هذا الأفق المعرفي، سعيا منه لإحلال الدراسة التراثية في موقع متميز من المقاربات اللسانية الحاسوبية، لاسيما الخليلية منها، فقد أدرك "أن ما تركه النحاة العرب الأولون شيء عظيم، ومفيد لاسيما بالنسبة للسانيات الحاسوبية التي تعتمد على الرياضيات والمنطق الرياضي"⁹.

وفي ظل هذا المسعى، حاول الدكتور عبد الرحمن الحاج صالح تجاوز مرحلة التردد لمقولات الخليل، والقفز بها إلى جملة التصورات التحليلية التي أفرزتها اللسانيات الحاسوبية، وفق رؤية تكاملية تتفاعل فيها الأطروحات التراثية مع الإفرازات الحداثية، وتتفاعل فيها مختلف الحقول المعرفية لتؤسس لنمط معرفي لساني، يتعامل مع اللغة بمنطق مغايرينأى عن مجرد "وصف وقائعها المنجزة، ومحاولة تفسير ظواهرها المتعددة بتعدد مستوياتها الصوتية والصرفية والدلالية والتركيبية"¹⁰، إذ يترهن إلى إلزامية التكامل المعرفي "المبني على التفاعل المتعدد بين العلوم، وعلى ضرورة العمل على [...] إدراك أنماط اشتغالها وفق معالجات متقدمة ومتعددة الاختصاصات"¹¹.

مشروع الذخيرة اللغوية عند عبد الرحمن حاج صالح :

لا نجانب الصواب، إذا أقررنا بأن مشروع حوسبة اللغة عموما، عرف منذ بواكيره الأولى في سبعينيات القرن الماضي وثبات مهمة، بدءا من مراحلها الجنينية بوصفه مشروعا للذكاء الاصطناعي لمعالجة اللغات الطبيعية، الذي عني في مخططاته الخوارزمية بمسألة التحليل الدلالي للمصفوفات المعجمية من خلال البرامج الرقمية التي عرضها كل من "لونار Lunar وشردلو Shrdlu". ولعل أهم إنجاز يُحسب لهذه

الانطلاقة الجريئة هي إيجاد الأرضية المناسبة للمعالجة اللغوية بوصفها كلية معلوماتية على غرار باقي الكليات (الضوئية، الصوتية)، حيث صممت المقابلات الرمزية الأولى لخاصية الحرف التي ترتبها إلى آلية النظام الثنائي اعتمادا على مبدأ القسمة الإقليدية للمقابلات المعلوماتية [1.0]. ومن هنا، كان من الطبيعي أن تُوجه كل الجهود اللاحقة إلى تطوير البرامج التي تعاملت مع الوحدة المعجمية من حيث هي "أساس مقطعي Raison segmentale" حيث أوجدت العمليات الحاسوبية للوحدة المعجمية خاصية رقمية مقابل الوظائف التي تؤديها المفردة في اللغة، وهو ما أسهم في بروز مجموعة من البرمجيات التي عُنيت بالمعالجات الرقمية للبنى الصرفية، وتشكلات الحقول الدلالية، وصولا إلى الترجمة الآلية للغات الطبيعية.

وفي ضوء هذا التعاقب المرحلي، بدأ التوق إلى مباشرة البرمجيات التي تعنى بالعمليات المنطقية للغة [البنى التركيبية les structures syntaxiques]، حيث ظهرت جملة من البرامج التي أعادت النظر في المقترحات الأولى التي ذهبت إلى أن منطلق المعالجة الرقمية اللغوية لا بد أن يؤسس له من منظور شكلاني Formel بدءا بالبنية التركيبية للغة، في نحو ما كان ينادي به الأمريكي (1953) Bar-Hillel¹²، إلا أن هذا التوجه لم يلق الاهتمام المرجو، إلا بعد بروز تلك التحولات الكبرى التي قدمتها النظرية التوليدية لنجوم تشومسكي من خلال المقابلات السُلمية التشجيرية لتمثالات البنى النحوية والعمليات المنطقية للغة. وهاهنا، انبرت البرامج الحاسوبية إلى تقديم نماذج رقمية des simulations تحاكي الأنموذج اللغوي الطبيعي.

وتساوقا وهذا التحول، جاء تصور الدكتور عبد الرحمن حاج صالح "الذخيرة اللغوية" ليستثمر في جملة من البرامج الحاسوبية التي وفقت في التطبيق على الأنموذج اللغوي العربي، بعيدا عن البرامج التي اشتغلت على المستويات اللسانية بشكل منعزل، حيث قدم مشروعه الطموح الذي انماز بتكامله وإلمامه بأكثر خصوصيات اللغة العربية، ويمكن أن نحصر النماذج الرقمية التي كان يرومها المشروع في الآتي:

- بنك معلومات آلي
- الذخيرة كمصدر لمختلف المعاجم والدراسات:
- المعجم الآلي الجامع لألفاظ العربية المستعملة
- المعجم الآلي للمصطلحات العلمية والتقنية المستعملة بالفعل
- المعجم التاريخي للغة العربية.
- معجم الألفاظ الحضارية (القديمة والحديثة)
- معجم الأعلام الجغرافية
- معجم الألفاظ الدخيلة والمولدة.
- معجم الألفاظ المتجانسة والمترادفة والمشتركة والأضداد¹³

ويتوضح من خلال صنافه البرامج الرقمية التي ضمّها مشروع الذخيرة، أن مسعى المعالجة الرقمية فيها لا يقف عند حدود إيجاد المقابلات اللفظية للصيغ الإفرادية على نحو الوظيفة التي تؤديها القواميس والمعاجم اللغوية، وإنما الغاية جاءت لتطاول الدراسات اللغوية المقارنة، أبحاث الفيلولوجيا الجامعة، وقد حدد ملامحها الحاج عبد الرحمن صالح في الآتي:

دراسة تطور معاني الكلمات عبر العصور، ودراسة ترددتها بالنسبة لعصر واحد أو مؤلف واحد، دراسة تردد المواد الأصلية وأوزانها في كتاب واحد أو عدة كتب دراسة صيغ الجمل بحسب الأغراض والموضوعات. دراسة أساليب الكتاب في كل عصر ودراسة اتساع رقعة الاستعمال للمصطلحات في عصرنا هذا. دراسة الأصوات العربية (من خلال الذخيرة الآلية الصائتة). دراسة مجالات المفاهيم الحضارية أو العلمية خاصة. تكشف المترادف والمشتراك من الألفاظ في الاستعمال في وقت معين. تكشف الغريب والشواذ إفراداً وتركيباً كيفاً وكماً وبالنسبة إلى كل مؤلف أو نص وكل عصر. دراسة صيغ الجمل وظواهر الفصل والوصل في الخطاب، ودراسات في المجاز والاستعارة والكناية، وغيرها من الصور البيانية.¹⁴

ويتبدى من الأهداف التي توخاها من مشروعه، انها كانت تنهض عن دراية كافية بخصائص اللغة العربية وعن جملة التباينات والفوارق الحاصلة في الشكل والوظيفة مقارنة باللغات السامية الأخرى التي خطت خطى عملاقة في المعالجة الآلية للغة، وذلك بدءاً من تقليده من شأن التركيز في تصميم تلك الخطاطات الرقمية للمقابلات اللفظية للغة العربية، وذلك نظراً للتحويلات الكبرى التي عرفتها الأخيرة في الاستعمال، ونظراً للاختلافات اللهجية التي طرأت على منطوقها، وكذا الخاصية الأسلوبية التي تميزها. ومن هنا، فقد جاءت دعوته إلى جمع أكبر رصيد لغوي ممكن والتحضير لمدونة رقمية شاملة تحتفظ على كل الأنساق والأسيقة التي وردت فيها، ومن ثمة ينتقل البحث في عناصر اللغة بشكل مسحي، تراعى فيه التطورات الفيلولوجية التي عرفتها ذات العناصر والأنساق.

وبالتمعن في طبيعة الغايات التي انتهى إليها الحاج صالح في مشروعه بوصفها نماذجاً تمتاح من إفراسات التحليل الآلي اللساني، فإننا ندرك بأن التقاطع الذي نلمح صداه على صعيد الطرائق الإجرائية التي أطرت لمنظومته الرقمية، تأتت مشروعيته في خضم التغييب القسري للبرمجيات التي عنيت بخاصية الحرف العربي وتمثالاته الوظيفية والاكتفاء بتبني المقاربات النظرية الضيقة التي أفرزتها البرمجيات الغربية، متجاهلة الاختلاف البين في الشكل والوظيفة للغة العربية، كما قوضت البرامج الرقمية المقدمة من إمكانية المعالجة الحقيقية لكامل المدونة اللغوية العربية، وذلك أنها غضت الطرف على المكنوز التراثي لهذه اللغة، فوقفت عاجزة عن استكناه الجوهر المادي للغة العربية.

ومن خلال قراءة متأنية لتصورات مشروع أحمد صالح تدفع إلى وضع آلية تنفيذية تؤدي إلى تجسيد أهدافه وتحقيق مراميها، يمكن أن نجعلها في الخطاطة الآتية:

1- الجمع الألفاظ الأنساق اللغوية من المدونات العربية

الخطاب المكتوب : المستويات اللسانية بما يتوافق مع المعيارية العربية

2- التحليل

الخطاب المنطوق: معالجة الكلام مع مراعاة التبدلات اللهجية للغة العربية

3- الترجمة الآلية بما يتوافق مع النماذج الانطولوجية للغة

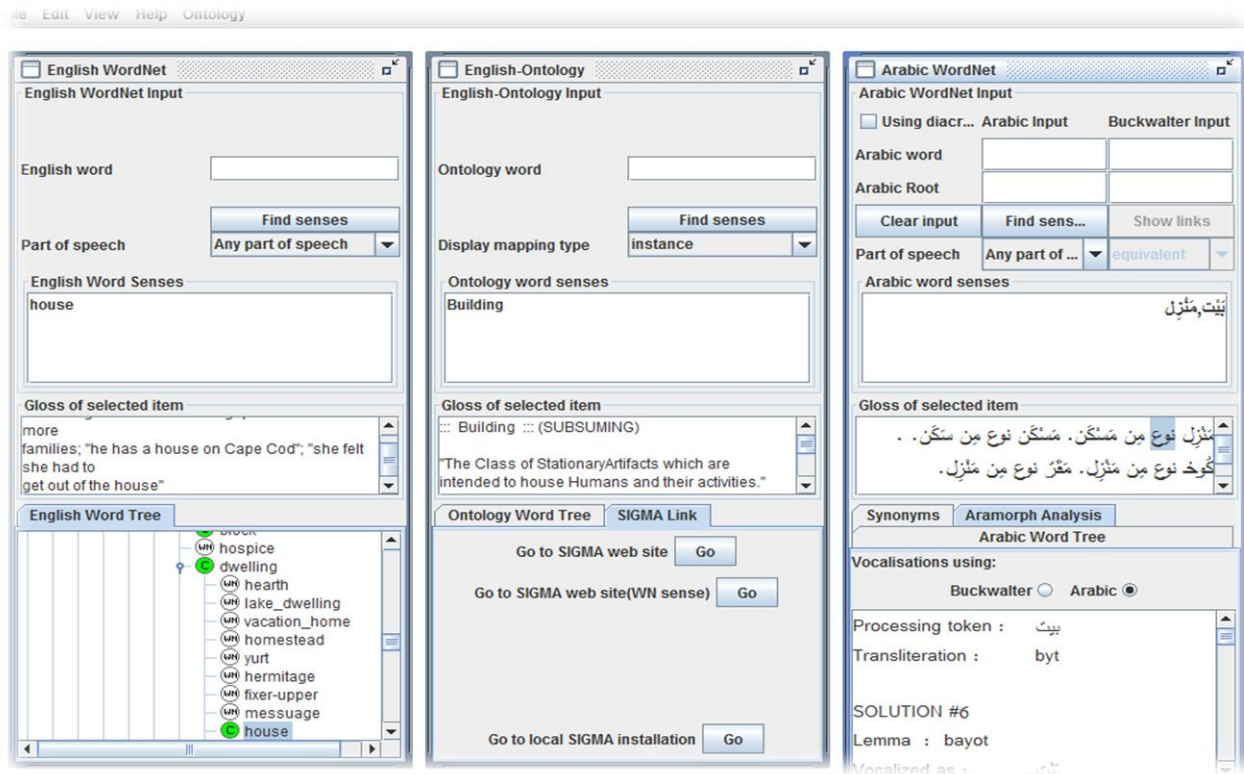
ولئن كانت الخطوة الأولى -الجمع- هي خطوة تستدعي إشراك كل المؤسسات والمراكز في الوطن العربي الكفيلة بتوفير مكنوزات المدونة العربية من نصوص لتصميم قاعدة بيانات كبيرة تؤدي إلى إعداد بنك معلومات (ArabiCorpus)، بوصفه أرضية معطيات كافية، لولوج الخطوتين التالية في المعالجة الآلية، فإن عملية التحليل هي أكثر العمليات تعقيدا، حيث يمكن توظيف كل البرامج التي يجتهد بعض الباحثين العرب على قلتهم في إصدارها في شكل: (kacst)، (Analyzer Buckwalter Arabic Morphological)، (Stemmer)، (kalmasoft)، والعمل على مراجعة نتائجها المقدمة، ومن ثمة تطويرها بطريقة تتشاكل مع طبيعة المعطيات المتجددة في بنك المعلومات.

الأنموذج التطبيقي المقترح للنظام المعالجة في الذخيرة اللغوية: (إعداد الدكتور ابراهيم بوداود والأستاذ مصمودي مجيد)

أرجأ الدكتور عبد الرحمن حاج صالح خطوة إقحام اللهجات العربية في آخر مشروعه، حيث قدم تصورا إجرائيا لكيفية معالجتها، معللا ذلك بجملة من الأهداف منها:

- تحديد القدر المشترك بين الفصحى ولهجاتها القديمة والحديثة
 - اكتشاف أسماء الحيوانات والنباتات في الأقاليم المختلفة
 - اكتشاف المصطلحات العفوية الحضارية والحرفية والصناعية والفلاحية وغيرها الجارية في اللهجات.
 - تحديد أوصاف النطق اللهجي ودراسة ظواهر الخفة في اللهجات
- وسعيا منا إلى تجسيد هذا المأخذ من المشروع، حاولنا الاستثمار في أحسن برنامج رقمي، يتيح لنا مكنة إدراج قاعدة بيانات لهجية للمنطوق الجزائري (غرب الجزائر)، حيث وقع الاختيار على برنامج WordNet بوصفه محاولة إنشاء قاعدة بيانات معجمية للغة الإنجليزية يستطيع المستخدم البحث من خلالها عن مفردة أو معنى وعلاقاته الدلالية مع مفردات أخرى وفق مجموعات. وقد انضمت إلى المشروع لغات أخرى كالعربية والفارسية حيث تم المحافظة على النموذج القياسي لـ WordNet الأوروبي EuroWordNet حيث تم تطوير البرنامج إلى اللغة العربية سنة 2015 الباحثان بوزوبع من جامعة محمد الخامس الرباط، المغرب، وبأولو روسو من جامعة بوليتكنيكا دي فالينسيا، إسبانيا (globalwordnet.org/arabic-wordnet/). النسخة قابلة للتحميل من الموقع :

<https://arabic-wordnet.soft112.com/download.html>



صورة لواجهة برنامج WordNet

يظهر على اليمين WordNet العربي، ويقابله على اليسار نطاق البرنامج WordNet الإنجليزي ويتوسطهما نطاق أنطولوجي باللغة الإنجليزية الذي يقوم بإنجاز العلاقة بين المفاهيم المراد البحث عنها باللغة العربية أو باللغة الإنجليزية.

وحيث يمكن استثمار هذه الإمكانيات في دراسة اللهجات المحلية العربية ومعرفة علاقتها بحفريتها التأثيلية من حيث الدلالة ومن حيث الهيئة النطقية وذلك من خلال تدعيم الوردنت العربي بقاعدة بيانات الذخيرة اللغوية العربية واستبدال النطاق الأيسر من البرنامج بإدراج قاعدة بيانات للهجات المحلية العربية، ثم ننشئ أنطولوجي للمفاهيم العربية في الواجهة الوسطى. ومن ثمة، عمدنا إلى إيجاد مقابلات رقمية لكل الفونيمات والألفونات للمنطوق الجزائري transcription phonétique، يمكن الترميز لخاصيتها الرقمية من منطلق أن "القواعد الفونولوجية لا تبلغ بالتمثيل حد الكفاية، ذلك أنه ستتعد إقامة الفرق مثلا بين صوت/ت/ وسو/ط/، وقد يبلغ اللبس حدًا لا يغني فيه استرفاد سائر المستويات"¹⁵ ومن ثمة، ننقل إلى إنشاء قاعدة بيانات المعجمية للمقابلات¹⁶:

hamza	ء	ا. إ. و. ي. ؤ. ع. هـ
ba'	b	ب
t ā'	t	ت
ta'	t	ث
zim	z	ز
hā'	h	ح
hā'	h	ج
dal	d	د
dāl	d	ذ
rā'	r	ر
zāy	z	ز
sin	s	س
jin	j	ي
sad	s	ص
dād	d	ض
tā'	t	ط
zā'	z	ظ
ayn	ʿ	ع
gayn	g	غ
fa'	f	ف
qaf	q	ق
kaf	k	ك
lam	l	ل
mim	m	م
nūn	n	ن
ha'	h	هـ
waw	w	و
ya'	y	ي
pā'	p	پ
gā'	g	گ
	a	فتحة مفتحة
	a	فتحة وسطية
	ā	فتحة مرققة
	ū	ضممة مفتحة
	u	ضممة وسطية
	ū	ضممة مرققة
	i	كسرة مفتحة
	i	كسرة وسطية
	i	كسرة مرققة
	ā	فتحة طويلة مفتحة
	ā	فتحة طويلة وسطية
	ā	فتحة طويلة مرققة
	ū	ضممة طويلة مفتحة
	ū	ضممة طويلة وسطية
	ū	ضممة طويلة مرققة
	i	كسرة طويلة مفتحة
	i	كسرة طويلة وسطية
	i	كسرة طويلة مرققة

المصطلح	الصفة	التلفظ	المقابل الفصح	منطقة الاستعمال
زَعْتَرُ	(اسم)	/zā'tar/	سعر	تلمسان
سَارِيحُ	(اسم)	/sārīḥ/	صهريج	تلمسان
مُوخ	(اسم)	/mūḥ/	مخ	تلمسان
عُوشُ	(اسم)	/ʿū/	عش	تلمسان
كَادُومُ	(اسم)	/kāḍūm/	قادوم	تلمسان
نَكَلُ	(فعل)	/nākāl/	نقل	تلمسان
كُمُخُ	(اسم)	/kūmh/	القمح	تلمسان
نَكَّشُ	(فعل)	/nākāḥ/	فلّح	تلمسان
تَيْفَلِيلِسُ	(اسم)	/tiflīlīs/	السنونو	تلمسان
بَالَى	(فعل)	/bālā/	حزم التبن بعد	تلمسان
مُمَبَّرُ	(اسم)	/mambār/	منبر	تلمسان

ومن ثمة يتمثل لنا النطاق الأيسر الذي أدرجت فيه قاعدة البيانات اللهجية لمنطقة من الغرب الجزائري على النحو الآتي:

ووفق هذا المنوال، فقد اشتغلنا على تصميم القاعدة من خلال بنك معلوماتي مزود بأكثر من 3400 صيغة لهجية منطوقة من الغرب الجزائري، التي يمكن تعميمها على كل الأقطار العربية، حيث يمكن استبدال النطاق الانجليزي بخانات المصطلحات اللهجية، وإدراج الأسيقة المتضمنة للمصطلح (انطولوجي) حيث وردت اللفظة، فيتسنى للمبحر القبض على كل المدلولات الفصيحة ذات العلاقة باللفظة في صيغتها اللهجية، شرط إدراج خاصية الفونيم الترميزية caractère بشكل موحد في كل النطاقات وفي كل قواعد البيانات بشكل موحد.

إن هذا المسعى التأسيسي الذي استلهه الدكتور عبد الرحمن حاج صالح دفع بنا إلى اعتماد المعالجة الرقمية العينية للمنطوق العربي، وسيدفع حتماً بالباحثين المهتمين بهذا الحقل إلى انتهاز المنهج التجريبي اعتماداً على المكنة التكنولوجية المتاحة بعيداً عن استنباطات التنظير السطحي، وذلك قصد احتواء المجال الحيوي للسان العربي المرتد في أغلبه إلى عصب المدونة اللغوية التراثية الصارمة، والتوثيق به إلى مسالك الدراسة الحاسوبية عبر مآخذ ترجيعية تستشرف آفاق التحليل اللساني انطلاقاً من المتحول إلى الثابت، وهي في مجملها مسالك تأخذ بإجرائية الوصف والمعاينة التزامنية لحيثيات البنية اللغوية المدروسة وغايتها ملامسة وبلوغ قصبة الممكّنات التفسيرية للغة بالإزتهان إلى ما تؤديه وتمنحه الإمكانيّة الرقمية .

إحالات

- 1- سعد عبد العزيز مصلوح.(2004). في النقد اللساني، دراسات ومناقشات في مسائل الخلاف، عالم الكتب، القاهرة، مصر، ط1.ص277.
- 2-محمد عبد المنعم حشيش، معالجة اللغة العربية بالحاسوب، مقال منشور بمجلة استخدام اللغة العربية بتقنية المعلومات، المجلد الأول، 1993 فاس، المغرب- ص40
- 3- عبد الرحمن الحاج صالح.(2007) النظرية الخليلية لحدثة - مفاهيمها الأساسية-، مركز البحث العلمي والتقني لتطوير اللغة العربية، الجزائر، العدد 04.ص51
- 4- عبد الرحمن الحاج صالح.(2007) النظرية الخليلية لحدثة - ص52
- 5- عبد الرحمن الحاج صالح.(2007) النظرية الخليلية لحدثة - ص52
- 6- عبد الرحمن الحاج صالح.(2007) النظرية الخليلية لحدثة - ص51
- 7- بوعناني مصطفى.(2013). الصوتات المعرفية والمسارات الذهنية للإنجاز اللغوي، عالم الكتب الحديث، إربد، الأردن، ص07
- 8- بوعناني مصطفى.(2013). الصوتات المعرفية والمسارات الذهنية للإنجاز اللغوي، ص09.
- 9- عبد الرحمن الحاج صالح.(2007) النظرية الخليلية لحدثة، ص50.
- 10- بوعناني مصطفى.(2013). الصوتات المعرفية والمسارات الذهنية للإنجاز اللغوي، ص11.
- 11- بوعناني مصطفى.(2013). الصوتات المعرفية والمسارات الذهنية للإنجاز اللغوي، ص11.
- 12 Retore C. (2005), Syntaxe et traitement automatique des langues, proposition a une contribution a un volume a Claude mailier ,université de Bordeaux .p:03
- 13- عبد الرحمن الحاج صالح.(2007) النظرية الخليلية لحدثة - ص35-37.
- 14- عبد الرحمن الحاج صالح.(2007) النظرية الخليلية لحدثة - ص37-39.
- 15- موسى نهاد (2000). العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر، عمان، الأردن، ص26

- 16- بوداود ابراهيمي. (2012). الكتابة الصوتية للهجات الغربية الجزائرية. المشروع الوطني للبحث (مؤلف جماعي) ، جامعة وهران، الجزائر. ص. 09.

Références

- 1- **Bouanani, M. (2013)**, Cognitive sounds and mental processes of linguistic realization, Alam El-koutoub El-hadith, Jordan.
- 2- **Boudaoud, B. (2000)**, phonetic transcription of the dialects of western Algeria, National Research Project, Algeria, University of Oran.
- 3- **Hadj Salah, A. (2007)**, Modern theory of El-khalil, Basic concepts, Scientific and Technical Research Center for the Development of the Arabic Language, vol4.
- 4- **Hadj Salah, A. (1999)**, The Language Ammunition Project and its Scientific and Applied Dimensions, Algeria Journal of Arabic Language, Algeria, Vol2.
- 5- **Hechich, M. A. (1993)**, Arabic language processing with computer literacy, Morocco, Fes, the use of the Arabic language with the computerization technique journal, vol1.
- 6- **Maslooh, S. A. (2004)**. In linguistic criticism, Studies and Acculturations in Controversy Problems, Alam El-koutoub, Ed:1, Cairo.
- 7- **Moussa, N. (2000)**, Arabic language, Towards a new characterization in the light of computer linguistics, Arab Institute for Studies and Publishing, Amman.
- 8- **Retoré, C. (2005)**, Syntax and automatic language processing, proposition to a contribution to a volume to Claude mesiller, France, Bordeaux university.
- 9- <http://globalwordnet.org/arabic-wordnet/>

المراجع باللغة العربية

- 1- بوداود ابراهيمي. الكتابة الصوتية للهجات الغربية الجزائرية. المشروع الوطني للبحث (مؤلف جماعي) ، جامعة وهران، الجزائر. (2012).
- 2- بوعناني مصطفى. الصوتيات المعرفية والمسارات الذهنية للإنجاز اللغوي، عالم الكتب الحديث، إربد، الأردن. (2013).
- 3- سعد عبد العزيز مصلوح. في النقد اللساني، دراسات وثقافات في مسائل الخلاف، عالم الكتب، القاهرة، مصر، ط1. (2004).
- 4- عبد الرحمن حاج صالح. مشروع الذخيرة اللغوية وأبعاده العلمية والتطبيقية، مجلة اللغة العربية، الجزائر، المجلد2. (1999)
- 5- عبد الرحمن الحاج صالح. النظرية الخليلية لحديثة - مفاهيمها الأساسية -، مركز البحث العلمي والتقني لتطوير اللغة العربية، الجزائر، العدد 04. (2007)
- 6- محمد عبد المنعم حشيش. معالجة اللغة العربية بالحاسوب، مقال منشور بمجلة استخدام اللغة العربية بتقنية المعلومات، المجلد الأول، فاس، المغرب (1993).
- 7- موسى نهاد. العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر، عمان، الأردن. (2000)
- 8- Retore C. (2005), Syntaxe et traitement automatique des langues, proposition a une contribution a un volume a Claude mailler ,université de Bordeaux .
- 9- <http://globalwordnet.org/arabic-wordnet/>