قراءة في المحللات التركيبية للجملة العربية

محمد قصبي
معهد الدراسات والأبحاث للتعريب- جامعة محمد الخامس، بالرباط –المغرب
تحت إشراف
د. عبد الحميد الجهاد
معهد الدراسات والأبحاث للتعريب- جامعة محمد الخامس، بالرباط –المغرب
mohkasbi86@gmail.com

تاريخ الارسال 2020/12/20.تاريخ القبول 2021/04/28 تاريخ النشر:2021/11/01

الملخص:

نسعى من خلال هذا المقال إلى تقديم أمرين؛ أولهما بيان أهمية برامج التحليل التركيبي في تصميم العديد من التطبيقات في معالجة اللغة العربية، وأهم المقاربات المعتمدة في هذا الميدان، والأمر الثاني نقدم فيه قراءة في مختلف المحللات التركيبية التي تم بناؤها أو تطويرها من أجل تحليل الجملة العربية، وقد خلصت هذه الدراسة إلى تحديد مواطن ضعف وقوة هذه المحللات التركيبية؛ ذلك أنها تولد أشجارا تقدم فكرة عن بنية الجملة دون النظر في الوظائف النحوية للجملة في اللغة العربية؛ وترتكب أخطاء متعددة على مستوى الإسقاطات للمركبات، كما أنها نماذج تحليلية تحاكي الأعمال الغربية ولا تراعي الخصوصية التي يتميز بها تركيب اللغة العربية وهو ما يستدعي السعي إلى بلورة محلل تركيبي يعتمد موارد لسانية تأخذ بعين الاعتبار خصائص اللغة العربية ومقوماتها التركيبية، ولعل هذا ما نسعى إلى بنائه في إطار المشروع الذي يتم إنجازه بمعهد الدراسات والأبحاث للتعريب.

الكلمات المفاتيح: التحليل التركيبي، المحلل التركيبي، المقاربة اللسانية، المقاربة الإحصائية، المقاربة الهجينة

Abstract:

Through this article, we seek to present two things; the first is to explain the importance of compositional analysis programs in designing many applications in the Arabic language processing, in addition to the most important approaches adopted in this field. In the second matter, we will present a reading in the various compositional analyzers that have been built or developed in order to analyze the Arabic sentence. This study resulted in identifying the points of weakness and strength of these synthetic analyzers. This is because they generate trees that provide an idea of the sentence structure without considering the grammatical functions of the sentence in the Arabic language. Multiple errors are committed at the level of compounds' projections. Moreover, they are analytical models that simulate western works and do not take into account the specificity that characterizes the structure of the Arabic language. This

certainly calls for seeking to develop a synthetic analyst that adopts linguistic resources which take into account the characteristics of the Arabic language and its structural components. Perhaps this is what we seek to build within the framework of the project being implemented by the Studies and Research Institute for Arabization.
Key words: Synthetic Analyst, syntax parser, Linguistic Approach, Statistical Approach, Hybrid Approach.

1. مقدمة

يعد التحليل التركيبي أحد المكونات الأساسية في المعالجة الآلية للغات الطبيعية، فهو يشكل خطوة أساسية لتصميم العديد من التطبيقات في معالجة اللغة الطبيعية، مثل: المدقق الإملائي والنحوي والتشكيل الآلي واسترجاع المعلومات والتوليد التلقائي للجمل والترجمة الآلية. ويتسم تركيب اللغة العربية بالغنى والتنوع والتعقيد نظرا للطبيعة البنيوية لمستويات اللغة العربية، إذ إنها مستويات تتداخل فيما بينها لتشكل نسق اللغة ونظامها الخاص الذي يجعلها تتفرد بمجموعة من الخصائص.

لذلك فقد كان ولا يزال يشكل محور اهتمام الباحثين والدارسين الذين ما انفكوا يقدمون مناويل حوسبية تمكن الآلة من فهم اللغة وتحليلها فإعادة إنتاجها. وقد أدّى هذا الاهتمام إلى تصميم العديد من المحللات التركيبية من قبيل: ستانفورد بارسر(stanford parser)، وبيركلي(berkely parser)، وبيكل (bikel)، والمحلل النحوي المعتمد على معجم الوظائف النحوية (arabic lfg rule-basic parser).

وعلى الرغم من تطور البحث في هذا المجال، إلا أن الساحة العلمية لا تزال تعرف قصورا كبيرا في بلورة محلل تركيبي يستطيع تحليل الأشكال المختلفة للنصوص العربية. وبدافع من هذه الحقيقة، فقد سعينا إلى تقديم قراءة في مختلف الأعمال التي اهتمت بالتحليل التركيبي وسعت إلى بلورة محلل تركيبي اعتمادا على مقاربات مختلفة ومتنوعة، وذلك قصد الاستفادة من المنجزات السابقة ومحاولة تجاوز ما شابها من نقص في إطار مشروعنا الذي يهدف إلى بناء موارد لسانية لبلورة محلل تركيبي للغة العربية.

2. إشكالية البحث

ينطلق هذا البحث من واقع الدراسات والأبحاث في مجال الذكاء الاصطناعي والمعالجة الآلية للغات الطبيعية، التي شهدت تطورا ملحوظا ومتناميا منذ النصف الثاني من القرن الماضي، هذا التطور الذي مس كل المستويات اللغوية تنظيرا وتطبيقا (صوت، صرف، تركيب، دلالة). إلا أن الملاحظ بالنسبة للغة العربية هو أنها لم تستفد بشكل كاف من هذا التطور العلمي؛ حيث ظلت الأعمال المقدمة تفتقر للأساس النظري الذي يراعي خصوصيات اللغة العربية وتداخل مستوياتها اللسانية، لذلك، فالبحث يهدف إلى الوقوف عند مختلف الأعمال التي اهتمت بحوسبة اللغة العربية عامة، والتحليل التركيبي خاصة. وستتضمن هذه الورقة ثلاثة أقسام؛ سنخصص القسم الأول لعرض مختلف التحديات والإشكالات التي تطرحها اللغة العربية من حيث خصائصها عامة والتركيبية خاصة، في حين

سنخصص القسم الثاني للتعريف بالتحليل التركيبي وعمل المحلل التركيبي، وأهمية برامج التحليل التركيبي في ميدان المعالجة الآلية، بينما في القسم الثالث سنقدم فيه قراءة في مختلف الأعمال المنجزة حول التحليل التركيبي مصنفة بحسب طبيعة المقاربات التي تعتمدها؛ إحصائية أو قواعدية، أو هجينة تجمع بين الإحصائي والقواعدي، وفي الختام قمنا بتسجيل أهم الاستنتاجات التي توصلنا إليها من خلال هذه القراءة، والتي من شأنها أن تساعدنا في بلورة محلل تركيبي للغة العربية اعتمادا على تطوير أحد هذه المحللات أو بناء محلل تركيبي يتميز عن هذه المحللات السابقة ويتجاوز مختلف الصعوبات التي وقفت عائقا أمام دقة هذه المحللات.

3. القسم الأول: خصائص اللغة العربية وقضايا التحليل الآلي للتركيب العربي

تعتبر اللغة العربية من بين اللغات الأكثر تعقيدا في علاقتها بالمعالجة الآلية، وذلك لما تتميز به من خصائص ترتبط بطبيعة بنائها الصوتي والصرفي والتركيبي والدلالي والتداولي، والمعجمي، وهو ما يجعلها تختلف عن اللغات اللاتينية، التي استطاعت أن تقطع أشواطا كبيرة في معالجتها آليا، والاستفادة من مختلف تقنيات التطبيق الآلي على مستويات تحليل أنظمتها الصوتية والصرفية والنحوية والدلالية... وهو الأمر الذي ما تزال اللغة العربية قاصرة على بلوغه؛ بما تتمتع به من خصائص ومميزات نذكر أهمها:

- **الطبيعة اللاصقة للكتابة العربية**(Agglutinante) ذلك أن نظام الكتابة في اللغة العربية يتميز بطابعه الخطي فنجد كلمات تختزل معلومات صرفية ونحوية ودلالية، مثل: كلمة **"أستذكرونهم؟"** تشكل جملة ذات بناء نحوي يعتمد:

حرف الاستفهام+ الفعل المضارع+ الفاعل(ض مت)+ المفعول به(ض مت)+ علامة الترقيم(؟)

- **الغنى والتنوع الدلالي الذي يسم المعجم العربي**، والناتج عن آليات الاشتقاق بأنواعها المختلفة، والاعتماد على الجذر في ترتيب مداخل هذا المعجم، فاللغة العربية لغة اشتقاقية تتعدد موادها ومداخلها المعجمية.
- **اللغة العربية لغة معربة** يتغير أواخر كلماتها بحسب مواقعها ووظائفها التركيبية، فمثلا تتغير أواخر كلمة "الأب" بحسب موقعها الإعرابي، نقول:

(1) أ- جاء الأبُ (الضمة علامة الرفع) الفاعلية

ب- رأيت الأبَ (الفتحة علامة النصب) المفعولية

ج- مررت بالأبِ (الكسرة علامة الجر)

كما تتغير هذه الحركات وتنوب عنها حروف، فمثلا الأسماء الخمسة، ترفع بالواو وتنصب بالألف وتجر بالياء، كما تتداخل هذه الحروف فتكون تارة دالة على الرفع كالألف في المثنى ودالة على النصب في الأسماء الخمسة، وتكون ظاهرة أو مقدرة...

- **التوسط اللغوي**: ويقصد به انحياز اللغة العربية نحو الشائع اللغوي، وتكره الشاذ والشارد، وتجمع بين كثير من الخصائص اللغوية المشتركة مع لغات أخرى.

- **الخاصية الصرفية**:(Marton, y. et al 2013) تتميز اللغة العربية باعتبارها من اللغات السامية بخاصية الاشتقاق الصرفي الجنسي على أنواع الصيغ، والأوزان، فأغلب أبوابها الصرفية تتميز بالاطراد.
- **المرونة النحوية**: حيث إن المقولات التركيبية تتمتع بالحرية في الانتقال بين التقديم والتأخير داخل الجملة العربية، مما يثمر أبنية جديدة للجمل العربية تتوزع بين الجملة الفعلية والجملة الاسمية،[1] خلافا لما تعرفه اللغات اللاتينية التي تتضمن بنية جملية قائمة على « SVO » عوض « VSO ».
- **الانتظام الصوتي**: يتسم النظام الصوتي للغة العربية بالبساطة، حيث تبتدئ المقاطع الصوتية بالصامت، ولا يمكن أن نجد مقطعا مؤلفا من صامتين متتاليين.
- **الحساسية السياقية**: تتميز اللغة العربية بحساسية سياقية عالية؛ إذ على مستوى الكتابة يتوقف شكل الحرف العربي على الحرفين السابق واللاحق له، أما على مستوى النحو فتتجلى في المطابقة بين الاسم الصفة تذكيرا وتأنيثا، تثنية وجمعا، وحركة إعراب...

إن هذه الخصائص2 وإن كانت تمثل مميزات للغة العربية فإنها في حد ذاتها تشكل عوائق وتحديات أمام المعالجة الآلية للغة العربية في مختلف مستوياتها، ونخص بالذكر المستوى التركيبي الذي يعد من أشد المستويات تعقيدا وأكثرها صعوبة أثناء الدراسة والتحليل؛ ذلك أن تحليل أية ظاهرة من ظواهر اللغة يستلزم تفاعل وتكامل مستوياتها الصوتية والصرفية من جهة والتركيبية والدلالية من جهة أخرى فضلا عن الدور الذي يقوم به المستوى المعجمي في عملية التحليل هاته، ويمكن أن نجمل بعض هذه التحديات التي يفرضها التحليل التركيبي، فيما يلي:

- **غياب التشكيل في اللغة العربية** (Al-Taani et al 2012) يؤدي إلى الالتباس والغموض، إذ إن كل حرف يأخذ الحركة التي تناسبه داخل بنية الكلمة، التي تتغير بتغير الكلمات السابقة واللاحقة لها داخل السياق اللغوي.
- **التطابق بين الفعل والفاعل عددا وجنسا**، في مثل:

(2) أ- "خرج التلاميذ"

ب- "التلاميذ خرجوا"

ففي الجملة (1-أ)، نلاحظ أن الفعل والفاعل لا يتطابقان من حيث العدد، مما يخلق مشكلا على مستوى التوليد الآلي للجمل؛ ذلك أن الفعل "خرج" جاء مفردا، بينما الفاعل جاء جمعا.

- **غياب علامات الترقيم المنتظمة**(Al-Taani et al 2012) يجعل من الصعب تحديد فواصل الجمل ويزيد من الغموض على مستويات مختلفة بما في ذلك التحليل النحوي.
- **اعتماد اللغة العربية على الإضمار**، فتحذف الفاعل، فيكون ضميرا مستترا وجوبا أو جوازا، وهو ما يعقد مهمة التحليل، اعتمادا على بناء الجملة.

4. القسم الثاني: التحليل التركيبي وعمل المحلل التركيبي

يتأسس التحليل التركيبي للغة العربية على دراسة المكونات التي تتشكّل منها الجملة، انطلاقا من تحديد بنية الجملة ووظائف مكوناتها ومواقعها كالفاعلية والمفعولية والمبتدأ والخبر...، وما فتأت الدراسات اللسانية تقدم نظريات تهدف إلى وصف وتفسير الآليات التي تتم بها عملية إنتاج اللغة، فمنذ الثورة اللسانية التي أحدثتها النظرية التوليدية التحويلية في الدرس اللساني وموضوع الجملة يأخذ الحيز الأوفر في بناء النظريات والتصورات الكفيلة بتفسير وتمثيل بنية الجملة باعتبارها مقولة ذهنية مجردة، لها حضور بالقوة في مستوى الكفاية، وحضور بالفعل على مستوى الأداء أو الإنجاز. وهما المنحنيات التي اضطربت فيهما النظرية وظلت تقدم العديد من النماذج بدءا من كتاب البنى التركيبية "1957" ثم ظهور النظرية المعيار، والمعيار الموسعة ما بين "1976" و "1977"، ثم بروز نظريات ومفاهيم الربط العاملي والمبادئ والوسائط "1980" و انتهاء بالبرنامج الأدنوي "1993".

ولعل الناظر في مختلف هذه النماذج النظرية يلحظ ذلك التوجه بالدرس اللساني نحو الصورنة والبحث في الآليات المعرفية والذهنية لتوليد الجمل، وفي هذا السياق نميز بين مقاربتين بارزتين هما:

- ✓ **المقاربة الأولى:** وتقوم على تحليل الجملة إلى مكوناتها المباشرة، وتسمى بالنحو المكوناتي (Constituency Grammar) الذي يضم اللسانيات التوليدية بمختلف مدارسها واتجاهاتها وتطوراتها من البنى النحوية إلى البرنامج الأدنوي مرورا بكل من نحو القواعد المعممة (Generalized Phrase Structure Grammar)، والنحو الوظيفي المعجمي (Lexical Functional Grammar)، والنحو الموجه بالرؤوس (Head Driven Phrase Structure Grammar) وتقوم هذه المقاربة على فرضية مفادها أن مكونات الجملة لا تؤدي وظائفها النحوية فرادى حتى تنتظم في مركبات تجتمع حول رأس واحد.
- ✓ **المقاربة الثانية:** تقوم هذه المقاربة على تحليل الجملة بناء على العلاقات التركيبية القائمة بين الكلمات، وتسمى بالنحو الاعتمادي أو العلائقي (dependency grammar)الذي تعود جذوره الأولى إلى أعمال بانيني (Panini) حول اللغة السنسكريتية، والتي قام بإحيائها اللغوي الفرنسي لوسيان تسنيير (lucien tesnière) من خلال كتابه النحو التركيبي.3 وتقوم على ثلاثة مفاهيم متكاملة، هي التحويل (transfer) والربط(connection) والتوزيع (junction).

وفي ظل هاتين المقاربتين نجد العديد من المحللات التركيبية التي هي عبارة عن برامج تقوم بتحليل بنية الجملة، وذلك من حيث ترتيب عناصرها والعلاقات التركيبية والوظيفية التي تربط بين عناصرها أو مكوناتها، ويستند في ذلك إلى مجموعة من الخطوات التحليلية التي تدمج تقنيات وأدوات تيسر عملية المعالجة الآلية. لذلك فالتحليل التركيبي له أهمية في بلورة العديد من التطبيقات كالتدقيق النحوي والتشكيل الآلي والأنظمة المساعدة في الترجمة وتركيب الكلام وتحليله، بالإضافة إلى عدد من التطبيقات في مجال استخلاص المعلومات والتنقيب في النصوص مثل نظام الإجابة الآلية عن الأسئلة والتلخيص الآلي واستخلاص المصطلحات.

5. القسم الثالث: نماذج المحللات التركيبية وتصنيفها

إن اهتمام المحلل التركيبي بالمستوى النحوي/التركيبي لبنية الجملة من حيث مكوناتها وعناصرها وعلاقاتها ووظائفها، يجعله مجالا خصبا لمختلف الدراسات والأبحاث التي تحاول توصيف هذا النظام النحوي الذي يتحكم في عناصر الجملة تقديما وتأخيرا، رفعا ونصبا وجرا... وتمتزج فيه متغيرات ترتبط تارة بما هو صوتي أو صرفي أو تركيبي أو دلالي أو تداولي مما يجعل مهمة هذه البرامج أعقد بكثير مقارنة مع اللغات الطبيعية الأخرى، لذلك فقد سعى العديد من الباحثين إلى بلورة محللات تركيبية للغة العربية تتوزع إلى محللات تعتمد المقاربة الإحصائية أو القواعدية أو الهجينة، وفيما يلي تقديم لأهم هذه الأعمال.

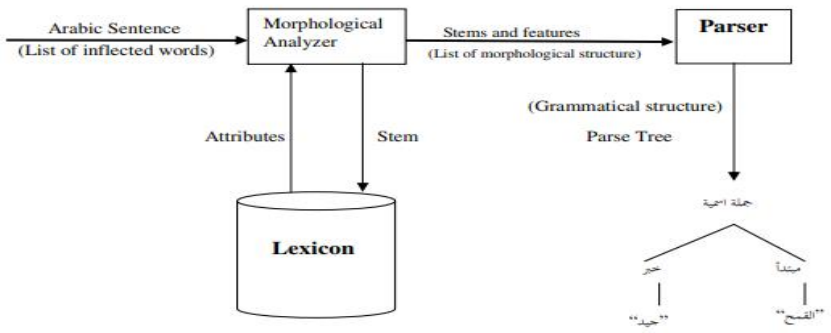
1.5. الأعمال القائمة على المقاربة اللسانية:

المقاربات اللسانية (Approches Linguistiques) تعتمد على جمع مختلف القواعد المبثوثة في كتب اللغة وإعادة توصيفها وصياغتها بالشكل الذي يمكن الحاسوب من معالجتها، وتشكل هذه القواعد، مادة البرامج التي تمكن من التحليل الآلي للظواهر اللغوية، بعد اختبار نسبة الدقة بها، واستكشاف مواطن الخلل وتداركه من خلال توصيف قواعد جديدة تختلف باختلاف أنواع الخلل.

والمقاربات اللسانية تعتمد على محاكاة منطق النظام اللغوي ومعرفة كيفية تمثله داخل العقل البشري، مما يمكنها من أن تكون منهجا تطبيقيا يعتمده الدارسون في الوصول إلى نتائج علمية ترتبط أساسا بالمستويات النحوية والدلالية والتداولية...، بالرغم مما تتطلبه من وقت طويل في بناء قواعد البيانات وتوصيف كل مستويات النظام اللغوي المعقد، وتباينه بين اللغات الطبيعية، الذي يستدعي تضافر الجهود وتكاثفها من أجل تحقيق الأهداف التي يطمح إليها الإنسان في علاقته بعصر التقنية والتكنولوجيا، ومن الأعمال التي اعتمدت هذا المنهج/ المقاربة، نذكر:

1.1.5. المحلل التركيبي للنص العلمي للغة العربية الحديثة المعتد على نحو الجمل المحددة

يعد هذا المحلل من المحللات التركيبية التي تم تطويرها باعتماد Prolog كأداة أساسية تستجيب لصورنة نحو الجمل المحددة (DCG)؛ ويستند إلى الخطوات التحليلية المبينة في الشكل(1):



الشكل(1) خطوات التحليل التي يعتمدها المحلل التركيبي

حيث يستخدم هذا المحلل التحليل الصرفي لمعالجة النص المدخل من خلال توفير مختلف المعلومات الصرفية للكلمات المكونة للجمل البسيطة والمركبة، وقد تم تزويد المحلل بمجموعة من القواعد التركيبية التي تمكنه من التعامل مع بعض مظاهر الغموض التي تنشأ عن بعض المظاهر الصرفية كالجنس والعدد؛ كما تم إغناء المحلل بمعجم، ويقدم نتائج التحليل في شكل شجري.

كما تم اختبار المحلل على 212 جملة فقط، استطاع المحلل تحليل 146 جملة بشكل صحيح ، وحوالي 44 جملة انتهت بتحليل غير صحيح 22 جملة لم يتم تحليلها، والجدول(1) يوضح نسب التحليل التي حققها هذا المحلل (شعلان وآخرون، 1999).

		Number of sentences	%
Parsable	**Grammatical Success**	146	68.8
	Grammatical Failure	44	20.7
Unparsable	**Ill-formed**	3	1.5
	Failure	19	9
Total		212	

الجدول(1) نسب التحليل التي حققها المحلل

2.1.5. المحلل المعتمد على قواعد التوحيد Unification Based Grammar

لقد طور (عثمان وآخرون 2003) محللا يقوم على رسومات مبيانية لتحليل الجملة العربية، اعتمادا على النحو الترابطي (FUG) الذي ينسجم مع بيئة نظام "برولوج" ويتكون هذا المحلل من 170 قاعدة مقسمة إلى 22 مجموعة كل منها تشكل فئة نحوية، ويحتوي هذا المحلل على محلل صرفي يتكون من شبكة انتقائية معززة بالعديد من السمات الصرفية والتركيبية إلى جانب اعتماد مدقق نحوي لحل مشكلة التدقيق النحوي للجمل العربية.

3.1.5. المحلل النحوي القائم على قواعد الارتباط (LG)

يعد هذا المحلل من المحللات التركيبية التي تعتمد المنهج القواعدي، المعتمد على النحو الترابطي (Link Grammar) الذي ترجع أعماله لكل من ديفي تيمبيرلي ودانييل سليتر ويقوم على بناء علاقات ثنائية بين الكلمات، حيث استفاد وارن كاسبير وغيره من الباحثين من الأعمال السابقة في تحليل اللغات الطبيعية باستخدام قواعد الارتباط، واستطاعوا تطوير هذا المحلل التركيبي للغة العربية الذي يعتمد على خطوتين تحليليتن، هما (كاسبير وآخرون، 2007):

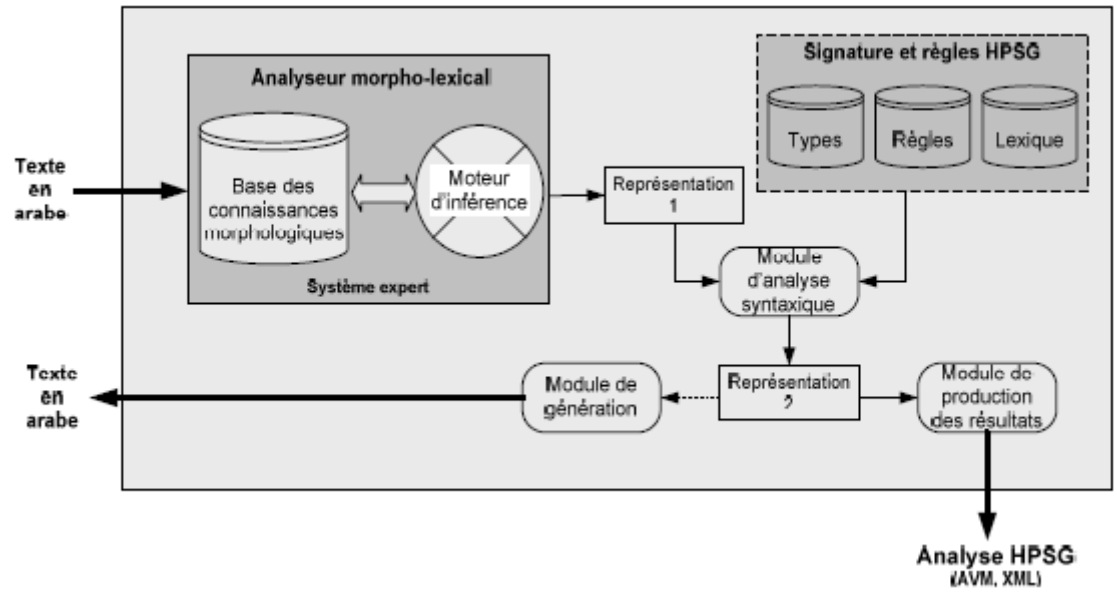
- **الخطوة الأولى** ترتبط بما قبل المعالجة وتهتم بالكتابة بالحروف اللاتينية، وتصنيف الكلمات حسب وظائفها بناءً على معجم يحتوي على متطلبات الربط لكل كلمة، بالإضافة إلى اعتماد محلل (Buckwalter) الصرفي.
- **الخطوة الثانية** ترتبط بالتحليل التركيبي اعتمادا على قواعد النحو الترابطي من أجل تحليل الجمل المدخلة في شكل شجري.

لقد تم تصميم هذا المحل ليحلل الجمل ذات النمط التركيبي (فعل+ فاعل+ مفعل به) من خلال تقديم مختلف التحليلات الممكنة وفقًا لقواعد النحو الترابطي، كما يستطيع هذا المحلل معالجة العديد من الظواهر المورفولوجية كالإضافة والنفي، والشرط والتقديم والتأخير بين مكونات الجمل.

والجدير بالذكر أن الأعمال المنشورة حول هذا المحلل تظهر العديد من المميزات والقدرات التحليلية له دون تقديم أية تفاصيل حول طبيعة القواعد المعتمدة كما أن نتائج اختباره والنسب التي حققها غير متاحة.

4.1.5. المنصة المتكاملة للتحليل النحوي (PHARAS)

لقد قدم الباحث مراد لوكام من جامعة حسيبة الجزائرية منصة تكاملية لتحليل النص العربي تحليلا نحويا يعتمد النحو الموجه بالرؤوس (**HPSG**)؛ حيث تعتمد هذه المنصة على خطوات تحليلية تبتدئ بتحليل النص المدخل وتقسيمه إلى كلمات، ثم تحليل هذه الأخير تحليلا صرفيا ومعجميا، وتركيبيا اعتمادا على القواعد التركيبية للنحو الشكلي (**HPSG**)، و(**AVM**)، للوصول إلى تحليل نحوي ودلالي للنص العربي المدخل، والشكل(2) يوضح خطوات التحليل التي يعتمدها هذا المحلل التركيبي (لوكام مراد ولسكري محمد 2008):



الشكل (2) يوضح خطوات التحليل التي تعتمدها منصة

ومن خلال اطلاعنا على الورقة البحثية لم يشر الباحثان إلى اي معلومات حول اداء النظام ونتائج اختباره.

5.1.5. المحلل التركيبي البسيط المعتمد على المنهج القواعدي

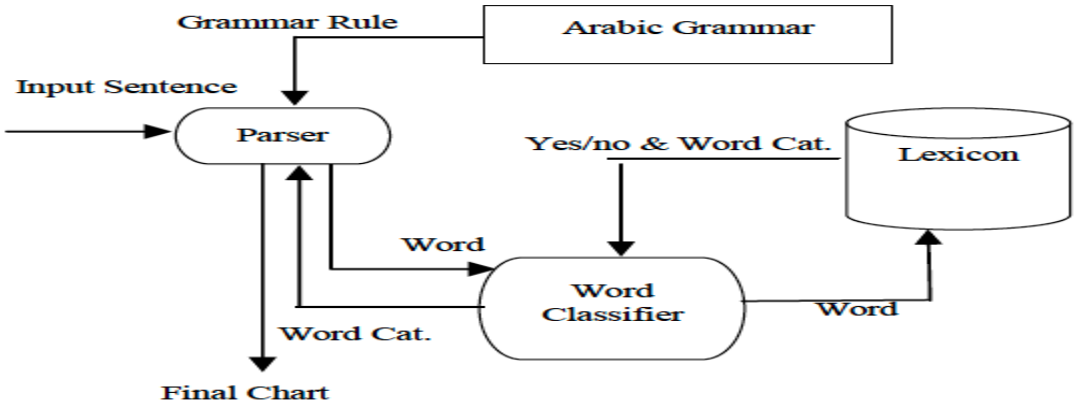
في سنة 2011 عمل الباحثان(محمد منى وعمر نزلي) على تطوير محلل تركيبي للغة العربية المقاربة المعتمدة على القواعد لتحليل الجملة العربية؛ حيث يقوم هذا المحلل على ثلاث مراحل، هي:

- **المرحلة الأولى:** وترتبط بما قبل المعالجة وتتضمن ثلاثة عناصر تحليلية اختيارية، هي normalization, tokenization, POS disambiguation
- **المرحلة الثانية:** بناء مدونة تتكون من 350 جملة تتضمن 9339 كلمة، تم توسيمها باعتماد برنامج (ASPOST POS tagger).
- **المرحلة الثالثة:** تمت بلورة العديد من القواعد التركيبية التي تسهل عملية التحليل التبعي للجملة العربية.

ويقوم هذا المحلل الذي طوره كل من منى علي محمد وعمر نزلي على اعتماد القواعد التركيبية لتحليل الجملة العربية بلغت 150 قاعدة تركيبية مشتقة من جمل المدونة المعتمدة، وقد تم اختبار المحلل على 70 جملة تحتوي على كلمات من مجالات مختلفة وحقق نتائج جد مشجعة بلغت حوالي 97٪ (محمد منى وعمر نزلي 2011)، إلا أنه لا يمكن اعتماده، نظرا لصغر حجم المدونة المعتمدة.

6.1.5. التحليل النحوي الجدولي للغة العربية الفصحى الحديثة المعتمد على نحو السياق الحر(CFG)

لقد طور الطعاني وآخرون محلا يعتمد التحليل النحوي الجدولي (**chart Parsing**)، ويحلل الجمل العربية من الأعلى إلى الأسفل(الطعاني وآخرون 2012)، كما يستند إلى نحو السياق الحر (**CFG**) في تمثيل قواعد الجملة العربية، ويقوم هذا المحلل على الخطوات التحليلية الموضحة في الشكل(3):



الشكل (3) خطوات التحليل التي يعتمدها المحلل التركيبي

ويحتوي هذا المحلل على ثلاثة مكونات أساسية، هي:

- **التوسيم (Tag)**: حيث يتم توسيم كلمات الجمل المدخلة وتصنيفها إلى أسماء، وأفعال، وأدوات، وظروف...
- **القواعد التركيبية** المعتمدة على نحو السياق الحر(**CFG**)، التي تستهدف تحديد نوع الجملة المدخلة اسمية أو فعلية، اعتمادا على أقسام الكلام (POS).
- **خوارزميات التحليل النحوي الجدولي من الأعلى إلى الأسفل** من أجل تمثيل الجملة المدخلة في شكل شجري يرمز لرأسها بالرمز(S).

وقد تم اختبار المحلل(الطعاني وآخرون 2012) على 36 جملة اسمية حقق المحلل دقة بلغت 97.2٪، بينما لم يتجاوز المحلل نسبة 91.2٪ لما تم اختباره على 34 جملة فعلية.

2.5. المحللات التركيبية القائمة على المقاربة الإحصائية:

إنه بالرغم من أهمية المقاربات اللسانية وأجرأتها في تقديم برامج تمكن الحاسوب من معالجة اللغات الطبيعية، تظل تشكل حملا ثقيلا على الدارسين، لما تتطلبه من نفس طويل في سبر أغوار النظام أو الأنظمة اللغوية للغات الطبيعية، وتوصيفها التوصيف المحكم الذي يأخذ بعين الاعتبار مختلف الوضعيات التخاطبية والتداولية لكل اللغات الطبيعية، ونسج العلاقات فيما بينها. الأمر الذي يدفع العديد من هؤلاء الدارسين لاعتماد المقاربات الإحصائية وانتقاء المواضيع ذات الارتباط الوثيق بهذه المقاربات من أجل بلورة محللات تركيبية تنطلق من بناء نموذج إحصائي، يهدف إلى تحديد الاحتمالات الإحصائية للتحليلات المختلفة للجمل، بناء على عدد ضخم من النصوص الموسومة بمعلومات لغوية كعينات للتدريب، كما يتخذ تقنيات إحصائية متعددة خاصة كتلك التي تعتمد على الأساليب الاحتمالية وينطلق من خطوتين أساسيتين، هما:

- **التعلم أو التدريب**: حيث يتعلم النظام المعرفة المطلوبة من المدونة التي تحقق الإشباع.
- **التحليل**: تشكل هذه الخطوة اختبارا حقيقيا للنظام لاستخدام المعرفة المكتسبة في الخطوة السابقة لتحليل الجمل وإعطاء التحليلات التركيبية للجمل المدخلة.

وفي هذا الصدد تم اعتماد العديد من الموارد اللسانية كـ(PATB) و(CATiB) و (PADT) والتي شكل ظهورها قفزة نوعية في مجال التحليل التركيبي المعتمد على المقاربة الإحصائية؛ حيث تم اعتماد هذه الموارد في بلورة العديد من المحللات من قبيل ستانفورد بارسر(stanford parser)، وبيكل (bikel). وبيركلي من جامعة كاليفورنيا محلل بيركلي(berkely parser)؛ حيث يمكن لهذا المحلل تعلم القواعد النحوية الأخرى من شجرة الأشجار، وقد استطاعت هذه المحللات تقديم نتائج جيدة في تحليلها للغة العربية، والجدول(2)[4] يوضح مستويات الدقة التي حققتها هذه المحللات:

Parser	Stanford	Bikel	Berkeley
Accuracy	0.802	0.775	0.831

الجدول(2): نسبة دقة التحليل للمحللات الثلاثة

ومن الأعمال التي تم إنجازها، نذكر:

1.2.5. المحلل الإحصائي المعتمد على موارد LFG:

لقد قدمت (لمياء تونسي وآخرون 2009) محللا إحصائيا يستخدم البنك الشجري المعتمد على موارد LFG من أجل تحليل الجملة العربية. كما استخدموا خوارزمية التوسيم (A3) لإثراء نتائجها؛ حيث حقق المحلل نتائج مشجعة بلغت 77.78%.

هذا وقد اقترح الباحثون مرحلتين لتحسين أداء المحلل؛ تنبني المرحلة الأولى على التحليل باستخدام النحو الأكثر احتمالا لتحديد بنية الجملة المراد تحليلها، بينما في المرحلة الثانية يتم تدريب النظام على تحديد الوظائف باعتماد بنك المشجرات العربي ATB لتشكل مدخلات لخوارزمية التوسيم (A3).

2.2.5. محلل ستانفورد بارسر stanford parser

يعد محلل ستانفورد بارسر stanford parser من بين المحللات النحوية التي تعتمد الطريقة الإحصائية في معالجة اللغة الطبيعية، وقد صممته مجموعة ستانفورد لمعالجة اللغة الطبيعية، ويستخدم هذا المحلل البيانات المكتوبة لعدة لغات كالإنجليزية، والألمانية، والعربية والصينية، وقد تم تطويره وصيانته منذ سنة 2002 على يد Dan Klein وChristopher Manning.

كما أنه يعد من البرامج المفتوحة المصدر، والقابلة للتعديل والتطوير بناء على مقتضيات بنود رخصة جنو العمومية (GNU GPL). وهو متوفر على موقع جامعة ستانفورد أو Stanford NLP Group ويتم تثبيته اعتمادا على JAVA 1.5 سواء في الحواسيب التي تعتمد نظام التشغيل ويندوز أو إينيكس/ لينيكس أو ماكوس.

إن اعتماد هذا المحلل على الطريقة الإحصائية يمكنه من استخدام البيانات كمدخلات؛ فبالنسبة للغة العربية يتم اعتماد بنك المشجرات العربي (ATB) في نسخه الثلاثة من أجل توسيم أجزاء الكلام اعتمادا على PCFG/ Probabilistic context-Free Grammar، كما يتضمن مخرجات متعددة بناء على نوع التحليل

الذي يرغب فيه المستخدم، فعلى سبيل المثال الجملة :"هذا الرجل هو سعيد" يتم تمثيلها وفق الشكل (4)، اعتمادا على بنك المشجّرات العربي (Penn Arabic TreeBank):

```
Your query :  هذا الرجل هو سعيد.
Taggging : هذا /DT الرجل /DTNN هو /PRP سعيد /NNP ./PUNC
Parse :
(ROOT
  (FRAG
    (NP
      (NP (DT هذا ))
      (NP (DTNN الرجل )))
    (NP (PRP هو ))
    (NP (NNP سعيد ))
    (PUNC .)))
```

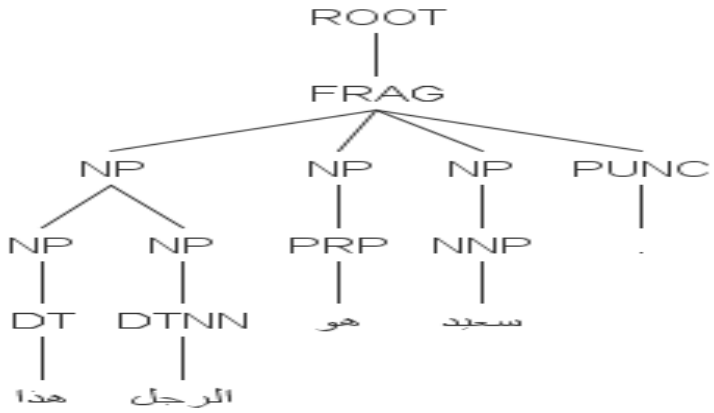
الشكل4: مخرجات المحلل ستانفورد بارسر لجملة "هذا الرجل هو

بينما اعتمادا على توسيم (PATB) [illegible]

```
(ROOT  (FRAG (NP (NP (DT هذا )) (NP (DTNN الرجل ))) (NP (PRP هو )) (NP (NNP سعيد )) (PUNC .)))
```

الشكل5: مخرجات المحلل ستانفورد بارسر لجملة "هذا الرجل هو

أو نتيجة تعتمد التمثيل [illegible]



الشكل 6: نتائج تحليل الجملة "هذا الرجل هو سعيد." وفق محلل ستانفورد بارسر

وتجدر الإشارة إلى أن المحلل ستانفورد في تحليله للتركيب العربي يعتمد العديد من الرموز التي أقرّتها شبكة البيانات اللغوية (LDC) بجامعة بنسلفانيا في عام 2003، التي يمكن أن نوضّح بعضها في الجدول التالي5:

الاختصار	المصطلح الإنجليزيّ	المصطلح العربيّ
JJ	Adjective	الصِّفَة
RB	Adverb	الظَّرف
CC	Coordinating Conjunction	حَرف عَطف
DT	Determiner / Demonstrative Pronoun	اسم إشارة
FW	Foreign Word	كلمَة أجنَبيَّة
NN	Common noun, Singular	نَكِرَة / شائع (مُفرَد)
NNS	Common Noun, Plural or Dual	نَكِرَة / شائع (مُثَنَّى أو جمع)
NNP	Proper Noun, Singular	اسم عَلَم (مُفرَد)
NNPS	Proper Noun, Plural or Dual	اسم عَلَم (مُثَنَّى أو جمع)
RP	Particle	أداة
VBP	Imperfect Verb (***nb: imperfect rather than present tense)	فعل مُضارِع / طَلَبيّ
VBN	Passive Verb (***nb: passive rather than past participle)	فعل مَبنيّ للمجهول
VBD	Perfect Verb (***nb: perfect rather than past tense)	فعل ماضٍ
UH	Interjection	أداة تَعَجُّب
PRP	Personal Pronoun	ضَمير شَخصيّ
PRP$	Possessive Personal Pronoun	ضَمير ملكيَّة
CD	Cardinal Number	عَدَد
IN	Subordinating Conjunction (FUNC_WORD) or Preposition (PREP)	أداة رَبط / حرف عطف
WP	Relative Pronoun	اسم مَوصُول
WRB	wh-Adverb	ظَرف بصيغَة الاستفهام

جدول(3): الرموز المعتمدة ومقابلاتها باللغة العربية

وبالرغم من أن نسبة الأداء التي يحققها هذا المحلل، والتي بلغت حوالي 85.5 بالمائة،6 إلا أنه لا ينجح في تمثيل العديد من أبنية الجمل في اللغة العربية.

3.2.5. المحلل الإحصائي (ARSYPAR)

لقد تمت بلورة هذا المحلل بعد النجاح الذي حققه مشروع MASPAR؛ حيث عمد (خوفي وآخرون 2013) من جامعة صفاقس إلى اعتماد طريقة التعلم الآلي لإنشاء محلل نحوي للغة العربية يعتمد على مرحلتين، هما:

- **مرحلة التعلم**: حيث تتم هذه المرحلة تحت الإشراف، ويتم فيها تدريب النظام على القواعد المستخرجة من المدونة باستخدام (SVM)، كما تم اعتماد الإصدر الثاني والثالث من بنك المشجرات العربي ATB.

- **مرحلة التحليل** تتم من خلال استخدام المعرفة التي اكتسبها النظام من مرحلة التعلم من أجل تحليل الجمل.

هذا وقد سعى الباحثون إلي تحسين أداء المحلل من خلال إضافة العديد من التوسيمات POS، والاعتماد على المحلل الصرفي مادا (MADA) الذي طوره (حبش وآخرون ،2013 ، 2009).

وبالرغم من النتائج المشجعة التي حققتها هذه الأعمال، إلا أنها لا تزال تعرف قصورا كبيرا في تحليل العديد من الظواهر التركيبية للغة العربية[7]، كما أنها تشترط وجود عدد ضخم من المدونات الموسومة صرفيا ونحويا، كما أن هذه الأعمال لا تسلم من العديد من الانتقادات خاصة من حيث المنهج الذي تتبعه، ويمكن إجمال أهم الانتقادات الموجهة للمقاربة الإحصائية فيما يأتي:

- أن نتائجها تضطرب باضطراب العينات اللغوية المعتمدة من حيث عدم شموليتها لمختلف الخصائص اللغوية المرتبطة بالظاهرة المدروسة.
- أنها تظل قاصرة عن الإحاطة بالمستويات التي تعرف نسبة كبيرة من الاحتمالات، كما هو الشأن في المستويات الدلالية والتداولية للغة العربية.
- ارتباط هذه المقاربات الإحصائية بنسبية التحليل اليدوي للعينات المعتمدة، وما يمكن أن يشوبها من قصور يؤثر بشكل مباشر في طبيعة النتائج المحصل عليها.

3.5. المحللات التركيبية القائمة على المقاربة الهجينة:

تنبني هذه الأعمال على اعتماد المقاربة اللسانية للتقليل من حجم الصعوبات التي يواجهها التحليل النحوي، ثم تعميم النتائج من خلال الدراسات الإحصائية، ويندرج ضمن هذه المقاربة الأعمال التي تشرف عليها جامعة ستانفورد في بلورة المحلل التركيبي اعتمادا على الموارد اللسانية المتاحة، والأعمال التي يقوم بها مختبر ميراكل بتونس لتطوير المحلل ستانفورد بارسر (Nadia Ghezaie Hammouda Kais Haddar 2017)

1.3.5 المقاربة الهجينة وتحسين المحلل التركيبي:

لقد عمل كل من كريستوفر مينينغ(Christopher Manning) وسبانس كرين(Spence Green) باعتماد المقاربة الهجينة من أجل تطوير محلل نحوي للغة العربية الفصحى الحديثة(MSA) وتحسين أداء المحلل التركيبي ستانفورد(Manning, C. and Green, S. 2010)، فاشتغل الباحثان وفق منهجية تهدف في بداية الأمر إلى تحديد مصادر الغموض في العديد من الظواهر التركيبية؛ كالإضافة والتركيب الاسمي والجمل، ثم العمل على تجاوزها من خلال بناء موارد لسانية.

وفي هذا السياق عمل الباحثان على إنشاء العديد من القواعد اللغوية الاحتمالية الخالية من السياق Probabilistic context-free grammar (PCFG)، معتمدين على توسيم يدوي يحدد العلاقة بين الكلمات المضافة والجمل الاسمية والضمائر، بالإضافة العمل على توسيم مدونة من الكلمات بأجزاء الكلام (POS) من أجل تحسين أداء المحلل التي بلغت نسبة 75.95 بالمائة عند اختباره على مدونة الاختبار، كما عمل الباحثان على دمج هذه الموارد اللسانية بالمحلل ستانفورد واختباره على النسخ الثلاثة لبنك المشجرات العربي (ATB) ومقارنة نتائجه بنتائج محللات بيركلي وبيكل، وقد خلص الباحثان إلى أن الموارد اللسانية لها دور كبير في تحسين أداء المحللات التركيبية، بالرغم من الصعوبات التي تعتري عملية بناء الموارد اللسانية ودرجة شموليتها لمختلف الظواهر التركيبية.

2.3.5. المحلل التركيبي الهجين للغة العربية

يندرج هذا المحلل في إطار المشروع الذي يتم إنجازه بجامعة مانشستر؛ حيث قام الباحثان Sardar Jaf و Allan Ramsay ببلورة محلل تركيبي للغة العربية (Tounsi, L et al 2009)، اعتمادا على خطوتين أساسيتين؛ تهتم الخطوة الأولى ببناء خوارزميات التعلم الآلي وإنشاء العديد من العلاقات التبعية بين الكلمات اعتمادا على النحو التبعي/ العلائقي (Dependency Grammar).

أما الخطوة الثانية فتنبني على سعي الباحثين إلى بناء محلل يقوم على تكامل قواعد النحو في ضوء العلاقات التركيبية بين الكلمات التي تم بناؤها في الخطوة السابقة.

وقد بدأ الباحثان عملهما بتحديد المصادر الرئيسة للالتباسات النحوية في اللغة العربية، والقضايا التي تجعل مهمة تحليل تركيب اللغة العربية مهمة صعبة، ثم قاموا ببناء خوارزميات تقيد التشكيلات المختلفة للكلمات داخل السياق التركيبي للجملة، وتساعد على تحديدها اعتمادا على قاعدة من البيانات النحوية كقسم الكلمة، وشكلها، وطولها، وموقعها الإعرابي، وعلاقتها بما قبلها وما بعدها...، ولا يعد التحليل تحليلا ناجحا إلى إذا استوفى كل هذه العناصر، ثم يتم الاحتفاظ بهذا التحليل الجزئي الناتج عن هذه الخوارزميات من أجل تحسين أداء المحلل.

وقد حقق الباحثان نجاحا كبيرا في بناء هذا المحلل خاصة على المستوى المنهجي فهو محلل قوي وفعال، بينما على مستوى الأداء لم يتجاوز نسبة 66 بالمائة، لذلك يسعى الباحثان إلى الرفع من هذه النسبة من خلال الزيادة من حجم البيانات النحوية المحددة في الخطوة الأولى.

6. المناقشة والاستنتاجات:

لا شك أن ما تم تقديمه يعد لبنة من لبنات السعي إلى تمكين اللغة العربية من التطور الحاصل في ميدان المعالجة الآلية للغات الطبيعية، وخطوة من خطوات تذليل مختلف العوائق والتحديات التي يفرضها تركيب اللغة العربية، بغض النظر عما شاب هذه الأعمال من نقص على المستويين المنهجي والتطبيقي؛ فقد لاحظنا أن جل المحللات

التي عرضناها الفقرات السابقة تعتمد المقاربة اللسانية في بنائها للمحلل التركيبي وذلك من خلال السعي إلى إنشاء قواعد تمكن من التعرف على بناء الجملة العربية؛ حيث واجهت هذه الأعمال العديد من المشاكل، نظرا للطبيعة البنيوية التي تتميز بها اللغة العربية وتداخل مستوياتها الصوتية والصرفية والتركيبية وهو ما فرض على الباحثين العمل على بلورة محللات تتجاوز جانبا من جوانب هذه المشاكل والتحديات فاتجه خالد شعلان إلى التركيز على حل مسألة الغموض بينما اتجه آخرون إلى التركيز على حل مشكل التركيب والعلاقة بين الكلمات داخل الجملة.

في حين اتجهت بعض الأعمال إلى اعتماد بيئات لغوية تحليلية تستند إلى القواعد، كاستخدام محمد عطية لبيئة اللغويات Xerox ، وبيئة NLTK التي اعتمدها باحثون من جامعة البلقاء التطبيقية، في حين وظف قيس حدار من جامعة صفاقس التونسية أداة بناء المعرفة اللغوية.

ومع التطور الذي شهدته المعالجة الآلية خاصة في معالجتها للغات الأجنبية ظهرت العديد من المدونات الموسومة تركيبيا وصرفيا، ونخص بالذكر بنك المشجرات العربي (ATB)[8] (Maamouri et al., 2004)، وكاتب(CATiB)[9] (Habash and Roth, 2009)، وبنك المشجرات الذي يقوم على النحو التبعي(PADT)[10] (Smrž et al., 2002; Hajic et al., 2004; Smrž et al., 2008)، والتي أعطت دفعة قوية في ميدان المعالجة الآلية بالرغم من قيامها على محاكاة النماذج الغربية ولا تحترم –كما يجب– طبيعة الخصائص التركيبية للغة العربية.

وفي الختام نستطيع أن نسجل أهم الملاحظات الاستنتاجات التالية:

- أغلب الأعمال المقدمة ترتبط بمدونات اختبار تعاني نقصا في تغطية العديد من الظواهر التركيبية للغة العربية، أو يتم معالجتها بشكل تجزيئي يهدف إلى تحسين أداء النظام في بيئة من البيئات.
- نقص خبرة اللغويين في مجال الذكاء الاصطناعي والبرمجيات الحاسوبية مما يفرض التعاون المتبادل بين اللغويين والحاسوبيين من أجل ردم الهوة بين هذين القطبين عوض الاشتغال في شكل جزر منفصل بعضها عن الآخر.
- إن كثرة الأعمال المنجزة تؤكد حجم المجهودات المبذولة من قبل الباحثين من أجل بلورة محلل تركيبي يكون أكثر دقة وقوة في تحليل مختلف الظواهر التركيبية للجملة العربية.
- أغلب الأعمال سعت إلى محاكاة المحللات التركيبية للغات الأجنبية مما تسبب في تعقيد مهمة التحليل واضطرابه سواء على مستوى القواعد أو على مستوى طبيعة التوسيمات النحوية المعتمدة مثلما نجد في المحلل التركيبي ستانفورد حيث إن عموم الرموز المعتمدة وعدم دقتها قد تكون وراء الخلط الذي يقع فيه المحلل بين بعض المكونات التي تسند إليها توسيمات مختلفة.

وفي ضوء ما سبق تظل الحاجة ماسة إلى بذل مزيد من الجهد من أجل بلورة محلل تركيبي يعتمد أدوات لسانية تراعي الخصائص التركيبية للغة العربية، بما تتضمنه من توسيمات تستجيب للخصائص الصرف تركيبية للمفردة اللغوية

واستعمالاتها سواء أكانت فعلا أو اسما أو حرفا. وتسهم في جعل المحلل التركيبي الذي نقترحه أكثر دقة وملاءمة للخصائص التركيبية للغة العربية، وهو الأمر الذي لا يزال يشغل اهتمامنا الفكري والعملي الذي نتبنى فيه تصورا نظريا يجمع بين تتبع تركيب اللغة العربية واستخلاص قواعده التركيبية، وفضيلة الانفتاح على مختلف المستجدات التي يعرفها ميدان المعالجة الآلية، وعلى رأس هذه الأدوات اللسانية، المعجم التركيبي والقواعد التركيبية للجملة العربية[11]، حيث قمنا ببناء معجم تركيبي يتضمن 24172 دخلة معجمية عملنا على توسيمها بسمات تركيبية بطريقة يدوية أو شبه آلية، تساعد على تحديد فئات الأفعال وانتقائها لمقولاتها، وتشكل عناوين تركيبية يتم استثمارها في خطوات التحليل التركيبي للجملة العربية. أما من حيث القواعد فقد استطعنا بناء 385 قاعدة تركيبية موزعة على أنواع الجملة العربية من فعلية واسمية وشرطية واستفهامية، بالإضافة إلى جملتي القسم والتعجب، كما ترتبط بأنماط تركيبية تحدد مكونات الجملة العربية وما يلحق بناءها التركيبي من أحكام وأحوال، وقد مكنتنا من إنشاء قاعدة من البيانات اللغوية التي ترتبط بتحديد الأنماط الأصول لتراكيب الجملة وما يتفرع عنها من صور وأبنية تركيبية.

7. خاتمة

لقد سعينا من خلال هذا المقال إلى تقديم قراءة في المحللات التركيبية للغة العربية، من أجل الوقوف على مواطن قوة وضعف مختلف هذه المحللات بما استطعنا الوصول إليه من مراجع، وسعيا إلى إبراز مدى حاجة الساحة العلمية لمحلل تركيبي للغة العربية، ينطلق من أدوات لسانية تحترم طبيعة اللغة العربية وخصائصها البنائية، وليس محاكاة للأعمال الغربية، وهو الأمر الذي ألفيناه في ثنايا هذا المقال، حيث تم اعتماد مقاربات متنوعة تنطلق تارة مما هو لساني وتارة أخرى مما هو إحصائي رياضي، وتارة ثالثة من مزج بين هذه وتلك من أجل الرفع من أداء تلك الأنظمة، وتحسين أدائها.

8. المراجع العربية

- نبيل علي، كتاب اللغة العربية والحاسوب، دار تعريب ، الكويت، 1988.
- المعتز بالله السعيد طه(2011) آلية التحليل التركيبي في اللغة العربية، اجتماع خبراء المحللات النحوية للغة العربية، 18-20 أبريل 2011. دمشق- سوريا.

9. المراجع الأجنبية

Al-Taani, M. M. Msallam, and S. A. Wedian, ―A top-down chart parser for analyzing arabic sentences‖, Int. Arab J. Inf. Technol., vol. 9, no. 2, pp. 109–116, 2012.

A. redjaimia et Nedezhda hilal Arabic Parsing Experience, June 2015.

Casbeer, W. Dehdari, J. and Lonsdale, D. 2007. A Link Grammar Parser for Arabic. Perspectives on Arabic Linguistics. Volume 20.

D. Cer, Marie-Catherine de Marneffe , Daniel Jurafsky, Christopher D. Manning. Parssing to stanford dependencies : Trade-offs betwen speed and accuracy. 2002.

Khoufi, N. Aloulou, C. and Belghith, L. 2013. ARSYPAR: A tool for parsing the Arabic language based on supervised learning. In proceedings of The International Arab Conference on Information Technology (ACIT'2013). Zarqa University, Jordan.
L. Tesnière, Eléments de syntaxe structurale. Librairie C. Klincksieck, 1959.
Loukam, M. and Laskri, M.T. 2008. PHARAS: Une plateforme d'analyse basée sur le formalisme HPSG pour l'arabe standard: Développements récents et perspectives. JED'08, Journées de l'Ecole Doctorale, 09/10 Juin 2008, Université Badji Mokhtar, Annaba, Algérie.
Marton, Y. Habash, N. and Rambow, O. 2013. Dependency parsing of Modern Standard Arabic with lexical and inflectional features. Computational Linguistics, 39(1), 161-194.
Manning, C. and Green, S. 2010. Better Arabic Parsing: Baselines, Evaluations, and Analysis. In Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010). Beijing, August 2010, 394-402.
Mohammed, M. and Omar, N. 2011. Rule Based Shallow Parser for Arabic Language. Journal of Computer Science 7, Science Publications.
Manning, C. and Green, S. 2010. Better Arabic Parsing: Baselines, Evaluations, and Analysis. In Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010). Beijing, August 2010, 394-402.
M. Maamouri and A. Bies. 2004. Developing an Arabic Treebank: Methods, Guidelines, Procedures, and Tools Workshop on Computational Approaches to Arabic Script-based Languages, COLING 2004.
Nadia Ghezaie Hammouda Kais Haddar, Parsing Arabic Nominal Sentences with Transducers to Annotate Corpora, Computación y Sistemas, Vol. 21, No. 4, 2017, pp. 647–656
Othman, E. Shaalan, K. and Rafea, A. 2003. A Chart Parser for Analyzing Modern Standard Arabic Sentence. The MT Summit IX Workshop on Machine Translation for Semitic Languages: Issues and Approaches. New Orleans, Louisiana, U.S.A.
Shaalan, K. Farouk, A. and Rafea, A. 1999. Towards An Arabic Parser for Modern Scientific Text. In Proceeding of the 2nd Conference on Language Engineering. Egyptian Society of Language Engineering (ELSE), Egypt, 103-114.
Smrž and J. Hajic. 2006. The Other Arabic Treebank: Prague Dependencies and Functions. In Ali Farghaly, editor, Arabic Computational Linguistics. CSLI Publications.
Tounsi, L. Attia, M. and Van Genabith, J. 2009. Parsing Arabic using treebank-based LFG resources. In proceedings of the LFG09 Conference. Stanford University, Stanford, CA 94305, United States, 583-586.
Youssef Tahir, Modélisation A Objets d'une Base de Données Morpho-Syntaxiques pour la Langue Arabe, Thèse de Diplôme de Doctorat, Spécialité Informatique 17 juin 2006.

[1]– يتميز تركيب اللغة العربية بالغنى والتنوع، ويسمح بتركيبات متعددة تتخذ فيها التراكيب اللغوية تموضعات متنوعة بشكل يجعلها تبدو غير قابلة للحصر، وهو ما يفرض وصفا لسانيا دقيقا اعتمادا على تصور نظري ينسجم مع منطلقات المعالجة الآلية.

[2]. ينظر نبيل علي، كتاب اللغة العربية والحاسوب، دار تعريب ، الكويت، 1988، ص 43– 67.

[3]- Lucien Tesnièr 1959

[4] – Abdellah redjaimia et Nedezhda hilal 2015

[5] – المعتز بالله. "2011"

[6]– Abdellah redjaimia et Nedezhda hilal "2015"

[7] – قدمنا في هذا الصدد ورقة تقييمية للمحلل ستانفورد، قيد النشر، في وقائع الندوة الدولية الثالثة بالمعهد 2019.

[8] - Arabic Treebank

[9] - Columbia Arabic Treebank

[10]- Prague Arabic Dependency Treebank

[11] - قدمنا في هذا السياق مقالا منشورا بالمجلة الدولية للسانيات العربية، المجلد السادس عدد 1-2/ 2020، ص269-310، بعنوان: "نحو بناء أدوات لسانية لبلورة محلل تركيبي للجملة العربية" أوضحنا فيها طبيعة هذه الأدوات اللسانية ومدى مساهمتها في بلورة محلل تركيبي للغة العربية، في إطار المشروع الذي يتم إنجازه بمعهد الدراسات والأبحاث للتعريب.