المعالجة الآلية للأفعال الخاصة: "كان وأخواتها أنموذجا"

ياسين الحركاوي، شيماء طاهر، حكيمة خمار
قسم اللغة العربية وآدابها، كلية الآداب والعلوم الإنسانية، جامعة محمد الخامس)، الرباط، المغرب
yassine_elharkaoui@um5.ac.ma
chaimae_tahir2@um5.ac.ma
hakima.khamar@flsh.um5.ac.ma

الملخص

تعد الكلمة الوحدة الأساسية الأولى لمعالجة الجمل آليا. وتنقسم الكلمة في اللسان العربي إلى كلمة تامة مستقلة معجميا، وكلمة خاصة تستقي استقلاليتها من وظيفتها السياقية. فالفعل منها مثلا، يرد ببنية صرفية وتركيبية تتشابه مع الفعل التام المتعدي، الشيء الذي يجعل الحاسوب ينتج جملا لاحنة آليا.

نسعى من خلال هذه الورقة إلى معالجة الكلمة الخاصة معالجة آلية تُمكن الحاسوب من التمييز بين الأفعال الخاصة والأفعال التامة المتعدية؛ على اعتبار أن الأفعال اللازمة تحتاج مركبا اسميا واحدا لإتمام معنى الجملة. على خلاف الأفعال التّامة المتعدية التي تحتاج إلى مركبين اسميين لإتمام المعنى المراد منه، وكذلك الشأن بالنسبة للفعل الخاص. وسنقتصر في هذه الدراسة على "كان وأخواتها" كنموذج للأفعال الخاصة.

الكلمات المفتاحية: المعالجة الآلية_ الفعل الخاص _ اللواحق-كان وأخواتها.

Automatic Processing of Arabic Stop Words: Kana and its Sisters as a Model

Abstract

A word is an essential unit for processing sentences automatically. A word, in the Arabic language, is divided to perfective word that is independent lexically and a stop word that can be independent in its contextual function, so the stop verb for example has a morpho-synthetic structure that is similar to the perfective transitive word, which makes computer produce automatic wrong sentences.

This paper aims to process stop words automatically to make the computer distinguish the difference between stop verb and perfective transitive verb; considering that the intransitive verb needs one nominal phrase to complete the sentence's meaning, but the perfective transitive verb needs two nominal phrases, as well the case for stop verb. This study is limited to ' kaana & sisters'.

***Keywords**: natural language processing _ Stop Verb _ Affixes_ kaana & sisters.*

1. المقدمة

أصبحت المعالجة الآلية للألسن الطبيعة من الحقول المعرفية الجديرة بالاهتمام، نظرًا للدور الذي تضطلع به في معالجة الظواهر اللسانية عبر تقنياتها الآلية الحاسوبية التي تخدم الألسن الطبيعية بشكل كبير في التنقيب والتلخيص والترجمة واسترجاع المعلومات. ورغم استفادة اللسان العربي من هذه التقنيات، إلا أن بعض الوحدات اللغوية العربية مازالت في حاجة إلى معالجة آلية، نخصص منها في هذه الدراسة الأفعال الخاصة: "كان وأخواتها أنموذجا".

نقصد بـ "الأفعال الخاصة stop verbs" هنا؛ تلك الأفعال التي تتميز بسمات مختلفة عن باقي الأفعال، وتكمن خصوصيتها في طبيعتها التركيبية، والنحوية، والصرفية والدلالية، وكذلك في التعرف الآلي عليها وفي المعالجة الآلية لها؛ إذ تعد أفعالا وظيفية تكمن أهميتها في وظيفتها السياقية. أما وظيفة هذه الأفعال غير السياقية، فتتمثل في غياب أدوار تميزها خارج السياق؛ مادامت قوتها تكمن في ورودها في سياقات معينة. وهي أيضا، أحد أقسام الكلمة الخاصة التي تنقسم إلى اسم خاص (أسماء الإشارة، والأسماء الموصولة...) وفعل خاص (كان وأخواتها...) وحرف[1] (حروف الجر وحروف النصب والجزم...). وتعد الكلمات الخاصة أيضا من الأدوات الأساسية للعديد من تطبيقات المعالجة النصية كاسترجاع المعلومات (Vegas, 2014, p. 1).

نروم من خلال معالجة كان وأخواتها آليًا وضع قواعد حاسوبية تسمح بالتمييز بين الأفعال الخاصة والأفعال التامة المتعدية، وقد عالجنا هذه الأفعال باعتماد منصة سَفَر (SAFAR) عبر عملية توليف هذه الأفعال بلواصق خاصة بالأفعال، فتبين من خلال عملية التوليف، إنتاج عدد هائل من التوليفات، بلغ (380271 توليفة) يقدر بعضها الصحيح بـــــ 124479 توليفة. أما البعض الآخر فلاحن، ويبلغ 255792 توليفة، فتبين أن منصة سفر لا تميز بين الأفعال الخاصة والتامة المتعدية، نظرا لطبيعتهما المتطابقة من الناحية الصرفية والنحوية والتركيبية. وبذلك، فهي تخضعهما لنفس القواعد، مما أدى إلى توليد توليفات لاحنة دلاليا.

يعد هذا الإشكال وإشكالات أخرى من بين التحديات التي يواجهها الباحثون في حقل المعالجة الآلية للسان العربي، نظرا لتعقيدات اللسان العربي من الناحية البنيوية والدلالية. وتتجلى الفجوات البحثية بهذا المجال في نقص الموارد اللسانية المتاحة مثل قاعدة البيانات الضخمة والمعاجم الشاملة، والتي تعتبر لبنة أساسية لتطوير وتحسين أنظمة معالجة الألسن الطبيعية بشكل عام واللسان العربي بشكل خاص. كما يشكل التعدد اللهجي، بدوره عائقا أمام توحيد معايير التحليل اللساني، مما يزيد من صعوبة تطبيق النماذج الآلية على نطاق واسع.

نهتم في ورقتنا هذه بمعالجة الأفعال الخاصة معالجة آلية بغية الإسهام في الإنتاج السليم للسان العربي وفهم أعمق وذكي للسياقات المجازية[2]. وسنعمل على معالجتها في محورين رئيسيين، حيث خصصنا المحور الأول لتعريف الكلمة التامة والخاصة وأقسامهما عند كل من النحاة القدامى واللسانيين، وانتقلنا في المحور الثاني إلى الإشكالات التي تعترض الآلة في التعرف الآلي على هذه الكلمات انطلاقا من طبيعته التركيبية والصرفية، لنفصل بعدها في أهم مراحل التحليل الحاسوبي للأفعال الخاصة.

2. الكلمة في اللسان العربي

تُعَدّ الكلمة في الألسن الطبيعية عنصرا أساسيا في تكوين جملة مفيدة، وتختلف أقسامها باختلاف الألسن الطبيعية ككل. ففي اللسان العربي تنقسم الكلمة إلى اسم وفعل وحرف، ويعد الفعل أحد أقسامها الهامة، نظرًا لطبيعة الرتبة التي بُنِيَ عليها اللسان العربي. وقد تعددت مفاهيم الكلمة بين النحويين القدماء واللسانيين المحدثين تبعًا لتعدد المستويات اللسانية التي تناولت الكلمة، انطلاقا من الخلفيات العلمية لكل مستوى وطرق تحديدها.

1.2. الكلمة التامة بين التراث والحداثة:

يعّد مفهوم الكلمة أحد الموضوعات المحورية في الدراسات اللسانية، وقد اختلفت تعريفاتها باختلاف المناهج التي تناولتها. ففي التراث النحوي العربي، قدّم سيبويه (ت.180هـ) في كتابه التأسيسي "الكتاب"، تقسيما للكلمة، مُعرّفا بأنها" اسم أو فعل أو حرف يؤدي معنى دون أن تكون اسمًا أو فعلًا" (قنبر، 1988، صفحة 12)، مما يعكس تركيزه على البعدين "النحوي" و"الوظيفي" في تحديد ماهيتها، بحيث تُحدد هوية كل منها من خلال دورها الإعرابي داخل الجملة ومعناها السياقي. وهكذا مثّل تحليل سيبويه نقلة نوعية في الربط بين البنية اللسانية ووظيفتها التعبيرية، جاعلا من الكلمة وحدة تركيبية قائمة بذاتها؛ وهنا نقصد الكلمة التّامة، لكنها أيضا متفاعلة مع السياق النحوي العام.

أما في المرحلة التالية، قدّم الزمخشري (ت.538هـ) في "شرح المفصل" تعريفا أكثر تحديدا لماهية الكلمة- واصفا إياها بأنها "اللفظة الدالة على معنى مفرد بالوضع، وهي جنس تحته ثلاثة

[1] لم نُسَمِّ الحرف بالخاص لأن معناه في غيره وبالتالي فإن جميع حروف المعاني خاصة وليس هناك قسم آخر تام كما في الفعل والاسم.

[2] انظر العربية والذكاء الاصطناعي، تحرير: المعتز بالله السعيد، الفصل الأول، المبحث الثالث.

أنواع: الاسم، والفعل، والحرف. والكلام هو المركب من كلمتين أسندت إحداهما إلى الأخرى" (يعيش، 2001، صفحة 70). هذا التعريف يبرز أهمية الوضع اللساني في تحديد معنى الكلمة، مما يعني أن الدلالة ليست اعتباطية، بل تستند إلى اتفاق لساني سابق. ويقف ابن يعيش عند هذا التعريف بعمق، موضحا أن اللفظة أعم من الكلمة، حيث تشمل الألفاظ المهملة والمستعملة. ويرى أن إضافة الزمخشري لعبارة "الدالة عن معنى" تهدف إلى استبعاد الألفاظ المهملة التي لا تحمل دلالة في ذاتها أو داخل سياق التي ترد فيه، بينما يوضح وصفه بـ"مفرد" أن الكلمة يجب ألا تكون مركبة، إذ إن التراكيب تحمل أكثر من معنى (يعيش، 2001).

أما مصطلح "بالوضع"، فيستخدمه الزمخشري للتمييز بين الألفاظ التي تدل على معنى بوضع لساني مقصود، وبين الأصوات الطبيعية غير المقصودة، مثل صوت "الشخير" الذي يصدره النائم التي تعبر عن حالته دون أن تكون موضوعة للسان. بناء على ذلك، يرى ابن يعيش أن الكلمة عند الزمخشري تلتزم بثلاثة شروط أساسية:

- الصوت اللساني: يقصد بها الأصوات المنطوقة.
- القصد والدلالة بالوضع: أي أن تكون ذات معنى متفق عليه لسانيا، وليس مجرد صوتي طبيعي.
- الاستقلال بدلالة محددة: يقصد بها تلك الكلمة التي تحمل معنى متكاملا بذاتها (ومقصود هنا الكلمة التّامة)، دون الحاجة إلى سياق تركيبي. أما في سياق الدراسات اللسانية المعاصرة، يقدّم عبد القادر الفاسي الفهري رؤية متعددة الأبعاد للكلمة، فيعتبرها "وحدة لسانية تتميز بملامح صرفية وتركيبية ودلالية ومعجمية وصِواتية" (الفهري، 1990، صفحة 37) ساعيا من خلال هذا التصور إلى بناء نموذج تحليلي شامل يُسهم في رصد بنية الكلمة العربية بشكل طبيعي، مع مراعاة التفاعل بين مستوياتها المختلفة. مما يبرز هذا التعدد في الأبعاد إلى التعقيد الذي يكتنف تحديد مفهوم الكلمة كما أشرنا آنفا، سواء في التراث النحوي أو في اللسانيات المعاصرة، حيث احتلت مسألة التعريف حيزا كبيرا نتيجة تباين المنطلقات المنهجية والأدوات التحليلية المعتمدة.

فإذا كان النحاة القدامى قد صاغوا مفهوم الكلمة انطلاقا من الأدوات التحليلية المتاحة لهم آنذاك، كـ"الوظيفة النحوية" و"الوضع اللساني"، فإن اللسانيين المحدثين وعلى رأسهم الفاسي الفهري، أعادوا تشكيل هذا المفهوم باستخدام أدوات تحليلية جديدة، كالأبعاد المذكورة آنفا، مع التركيز الخاص على "الكلمة العربية" باعتبارها الوحدة الأساسية في التحليل الآلي، سواء على المستوى الصرفي أو النحوي أو الدلالي. كما يُميّز في هذا السياق بين الكلمة التّامة (التي تحمل معنى مستقلا بذاتها حتى لو وردت خارج السياق)، والكلمة الخاصة (التي لا تكتسب دلالتها إلا ضمن السياق الذي ترد فيه، فإذا وُضعت منفصلة عن السياق، فإنها تفقد معناها وتصبح مجردة من الدلالة)، على عكس الكلمة التّامة.

2.2. الكلمة الخاصة " Stop Word"

يُقصد بالكلمات الخاصة تلك الكلمات التي ترد في سياق أو نص ولا تؤدي معنى أو وظيفة إلا في علاقتها بغيرها، فتلعب بذلك دورا مهمًّا في سياق معين. ومن ثم، فالأفعال الخاصة تختص بوظائف تركيبية وأدوار دلالية أثناء ورودها في سياق محدد.

لقد عالجت اللسانيات الحديثة الكلمة الخاصة من جوانب عدة، فأبرزت فراغها الدلالي في مقابل الكلمة التامة التي لا تحتاج إلى سياق لتحديد موضوعها وقيمتها. إن الحرف لا ينقسم إلى خاص وتام، وذلك لأن الحروف كاملة تتميز بوظيفتها التي لا تتبين إلا بإسنادها لسياق محدد. إلا أن المعالجة الحاسوبية تزود الكلمة الخاصة بنفس السمات المميزة للكلمة التامة، وذلك من قبيل الفعل (أكل = زمن + حدث، كان = زمن)، وقد تمت معالجة هذا الإشكال وإشكالات أخرى متعلقة بالكلمة الخاصة في منصة "SAFAR"[3] بناء على المحطات الأربعة الآتية، نأخذ منها على سبيل المثال معالجتها للحرف(Driss Namly, 2015, pp. 7-8):

1. **مرحلة المعالجة الأولية:** تنطلق هذه المرحلة بتقسيم الجملة المدخلة إلى وحدات لسانية منفصلة يتم فيها إزالة العلامات غير اللسانية (كالرموز وعلامات الترقيم)، حيث يتم في هذه المرحلة التعرف على طبيعة الكلمات المدخلة.
2. **مرحلة التحليل الصرفي:** يتم في هذه المرحلة فرز الكلمات المدخلة وتحديد خصائصها الصرفية وعلاقاتها وتأثيرها ببعضها البعض، فعلى سبيل المثال: "إِنَّ العِلْمَ نُورٌ" تجرد على صيغة (حرف نصب وتوكيد + مركب إسمي$_1$ + مركب إسمي$_2$)، والحاسوب يقوم بشكل الكلمات بناء على تأثير (إنَّ) في المركبات الاسمية التي بعدها باعتباره كاملا.
3. **مرحلة التحليل النحوي:** يتم في هذه المرحلة تصنيف الأدوات وفق وظيفتها بالسياق، بحيث يتخذ كل حرف مجموعته الخاصة (حروف الجر، حروف العطف ...إلخ). وهو ما يعمل على تحديد دور الحرف داخل الجملة، فالقالب النحوي للمثال "إِنَّ العِلْمَ نُورٌ" هو (حرف نصب وتوكيد + مركب إسمي$_1$ + مركب إسمي$_2$)، ما يعني أن المركبين الاسميين الذين تليها هما اسمها وخبرها.

[3] منصة سفر هي منصة مخصصة لمعالجة اللسان العربي نظرا للحاجة المتزايدة لتوفير أدوات قادرة على التعامل مع المحتوى العربي بالأنترنيت. تأسست هذه المنصة سنة 2008. https: //arabic.emi.ac.ma/alelm

4. مرحلة التحليل السياقي: تعتمد هذه المرحلة على كل المراحل السابقة، بحيث يتم تجميع نتائجها لفهم معنى الحرف داخل الجملة، فحرف الجر مثلا (في) يأتي للدلالة على المكان في المثال (كان عثمان في المدرسة).

5. مرحلة التحليل الآلي: بعد تحليل دلالة الحروف الواردة في الجملة وبنيتها ووظيفتها، يتم استخدام هذه المعلومات في تطبيقات لسانية مختلفة: كالترجمة الآلية، وتوليد النصوص، وتصحيح الأخطاء اللسانية. وتتجلى أهمية دقة المعالجة اللسانية في صحة

6. المعلومات المقدمة للمتصفح. فترجمة كلمة (but) إلى العربية، مثلا تترجم بناء على المراحل السابقة، مع الحرص على سلامة المعنى الأصلي للجملة عامة والحرف خاصة.

مما سبق، فالكلمة الخاصة هي كلمة سياقية لا معنى لها في نفسها بل لا قيمة لها إلا بربطها بسياق محدد. والمقصود بالسياق هنا تلك اللواصق الممكنة التي يمكن توليفها في بنية الكلمة الخاصة سواء أكانت سوابق أو مقحمات أو لواحق. وقد تختلف اللواصق باختلاف عناصر الكلم المعنية بالتوليف (الفعل -الاسم – الحرف)؛ فكان وأخواتها مثلا، هي أفعال تميزت بانفرادها عن الأفعال التامة التي تخلو من إشكالات عدة أهمها التعرف الآلي عليها ومعالجتها آليا.

3.الأفعال الخاصة Stop Verbs

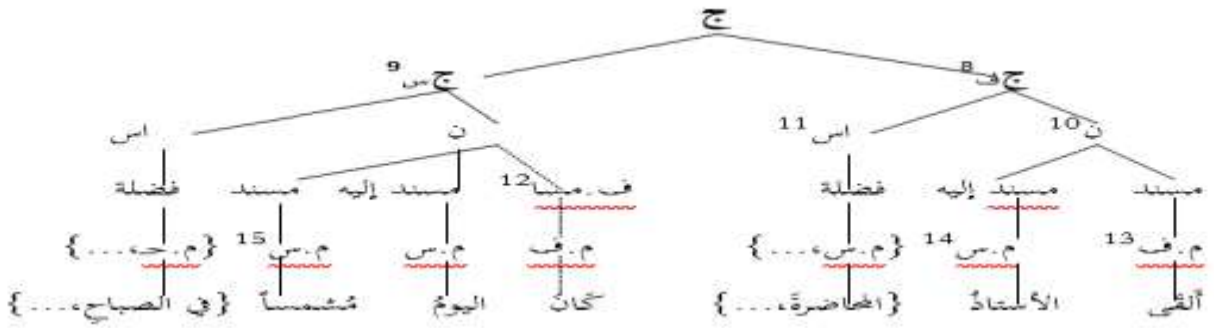
نقصد بالأفعال الخاصة تلك الأفعال ذات الوظائف التركيبية التي يتكرر استعمالها في جملة أو نص أو غيرهما. هذا الاستعمال يجعل هذه الأفعال تعمل في المقولات المعجمية وتؤثر فيها، وتتصدر الجملة العربية لكنها لا تتحدد وظائفها إلا داخل الجملة العربية؛ بمعنى أن وظيفتها لا يمكن أن تتحقق إلا إذا انتظمت داخل مقولات معجمية محددة (اسم1+اسم2). فهذه الأفعال الخاصة لا تُشكِّل مسندًا بذاتها، بل تحتاج إلى خبر لتفيد معنى تامًّا، عكس الفعل التام المتعدي الذي يُشكِّل مسندًا بذاته، ويتم بدون إسناده للفاعل الذي يكون اسما ظاهرًا أو ضميرًا فالاسم المنصوب (الخبر) هو الذي يقوم بدور المسند في الجملة مكان الفعل الخاص حتى يتمم المعنى عبر إشراكه في سمة الإسناد. وعند تجريد الفعل الخاص من السياق التركيبي (الجملة) يصبح المسند هو الخبر، وبالتالي فالفعل الخاص يفقد وظيفته عند إسقاطه من الجملة ولا يؤثر في تمام الجملة.

سميت هذه الأفعال بالخاصة لتجردها من دلالة الحدث، وهو ما يجعلها تختلف عن نظيرتها (أي الفعل التام المتعدي) في الكثير من الوظائف، نذكر منها صيغة البناء للمجهول التي لا تتوفر في هذه الأفعال لأنها ليس لها فاعل بل مبتدأ. أما من الناحية التصريفية، فلها خصوصيات تجعلها تنقسم إلى ثلاثة أنواع (متصرف تام-متصرف ناقص-جامد) ولكل نوع مميزاته التصريفية والدلالية والتركيبية في سياق الجملة العربية وكذلك الشأن في الاشتقاق. ورغم اختلاف مميزاتها وتعدد أقسامها إلا أن دورها موحد وهام في تأكيد المعنى العام للجملة العربية.

1.2.دور الأفعال الخاصة في اللسان العربي

تؤدي الأفعال الخاصة دورا وظيفيا بدخولها على الجملة الإسمية؛ إذ تدخل على المبتدإ والخبر فترفع الأول ويُسمّى اسمها وتنصب الثاني ويُسمّى خبرها، مثل: "كانَ البَرْدُ قارساً" (الدحداح، 1995، الصفحات 4-5). فالاسم المنصوب يقوم بدور المُسند في الجملة ويُتمِمُ المعنى النّاقص. وهذا يُخالف الأفعال التامّة المتعدية التي يكتمل معناها بذكر المرفوع ويكون المنصوب بعد ذلك فضلةً خارجة عن الإسناد. فالإسناد في الجملة الفعلية يخالف الإسناد في الجملة الإسمية نحو الشكل الآتي الذي يوضح بطريقة توليدية آلية الفرق بين الفعلين من الناحية البنيوية والسياقية:

خطاطة (1): أقسام الجملة العربية[4]



[4] ج ف: جملة فعلية. ج س: جملة إسمية. ن: نواة؛ أي مكونات رئيسية بالجملة. اس: استئناف؛ أي مكونات ثانوية بالجملة. ف.مسا: فعل مساعد م.ف: مركب فعلي. م.س:مركب إسمي.{...}: ما بين الحاضنتين يرتبط بنفس العلاقة الجدولية

يوضح الشكل أعلاه أقسام الجملة العربية التي تتفرع إلى قسمين؛ يختص أولهما بالجملة الفعلية التي تتكون من الناحية الإسنادية من شقين: الشق الأساسي والشق الثانوي أو التكميلي، كما يلي:

- الشق الأساسي: رمزنا له بـ"ن" التي تعني نواة أو عمدة الجملة، ويتفرع هذا الشق بدوره إلى فرعين متلازمين وهما (المسند والمسند إليه)، حيث تعود صفة **المسند** إلى **الفعل التام** و**المسند إليه** إلى **فاعل** أو **نائب الفاعل** في الجملة الفعلية.

- الشق الثانوي: رمزنا له بـ "اس"؛ ويعني استئناف (أي الشق الفرعي أو التكميلي) للجملة الفعلية حسب نوع الفعل[5] الوارد في الجملة من ناحية اللزوم والتّعدي.

في حين يختص الفعل الثاني بالجملة الإسمية التي تتكون من الناحية الإسنادية من شقين: الشق الأساسي والشق الثانوي:

- الشق الأساسي: رمزنا له أيضا بـ"ن" التي تعني نواة الجملة، تتفرع إلى فرعين متلازمين وهما المسند والمسند إليه، فالمسند إليه هو المبتدأ والمسند هو الخبر، وقد تدخل على هذه الجملة حروف ناسخة (إن وأخواتها) أو أفعال خاصة (كان وأخواتها)، وهي مكملات تتفرع عن النواة لتساهم في تأكيد الخبر وتقويته، وتتصدر الجملة الإسمية فتزيل الصدارة عن المبتدإ ويطلق عليه بعد هذا التحول (اسم كان). ونتيجة دخول هذه الأفعال على الجملة الإسمية، يتغير زمن الجملة وحركتا المبتدإ والخبر الإعرابية فقط.

- الشق الثانوي: هو الذي رمزنا له بـ "اس"؛ الذي يعني استئناف أي (الشق الفرعي)، وهو شق تكميلي في الجملة (فضلة) قد يرد نعتا أو جارًا ومجرورًا... أو غير ذلك.

وهكذا، يكمن الفرق بين الجملتين في طبيعة الفعل الذي يتصدرهما، فالفعل في الجملة الفعلية مركب أساسي يصل معناه إلى الفاعل والمفعول، وإذا ما تم إسقاطه يختل معنى الجملة وتختل مواقع المركبات التي بعده، نحو: "أَلْقَى الأُسْتَاذُ المُحَاضَرَةَ _ Ø الأُسْتَاذُ المُحَاضَرة"، فالجملة التي حُذف منها الفعل لا دلالة لها، أما في الجملة الإسمية التي يتصدرها فعل خاص لم تؤثر في دلالة وتركيب الجملة، نحو: "كَانَ اليَوْمُ مُشْمِسًا _ Ø اليَوْمُ مُشْمِسٌ"، لم تتغير الجملة لا دلاليا ولا تركيبيا، ولم تتغير أيضا بنية الجملة (مبتدأ وخبر). يتبين إذن أن الفعل التام المتعدي في الجملة الفعلية له دور أساسي وتربطه علاقة إسنادية تلازمية مع باقي مكونات الجملة، ويؤدي حذفه إلى لحنها. أما الفعل الخاص في الجملة الإسمية، لم يؤثر حذفه لا دلاليا ولا تركيبيا، كون دوره هو تغيير الزمن فقط وليس المعنى.

يتميز الفعل الخاص بمميزات عديدة أهمها أنه فعل له صيغة فعلية تشبه الفعل التام المتعدي في وزنه والأزمنة التي يتصرف فيها؛ إلا أن الاختلاف بينهما يكمن في عدم تضمن الفعل الخاص لحدث في نفسه. ويعتبر الحدث أهم مكون في الفعل التام المتعدي الذي يُمَكِّنه من التصريف التام دون نقص أي زمن وأيضا تصريفه التام بالماضي والمضارع والأمر رغم بعض الاستثناءات التي تتصرف تصرفًا ناقصا[6]. أما الفعل الخاص فيتصرف بعضه تصريفا ناقصا تارة بالماضي والمضارع فقط (نحو: مازال ومافتئ) أو جامدًا يلازم صيغة الماضي فقط (نحو: ليس ومادام). إلا أن أهم شروط عمله هي دخوله على الجملة الإسمية وليس الفعلية، إضافة إلى أن السياق الذي يرد فيه يتكون من {فـخ+ مب + خب}[7]. أما السياق الذي يرد فيه الفعل التام المتعدي فيتكون من {فـت + فا + مف}[8]. وهذا ما يجعل التوليف الآلي بين اللواصق والفعل الخاص يقدم احتمالات تؤدي في أحيان كثيرة إلى لحن في العديد من هذه التوليفات، مما يؤدي إلى إشكالات خلال معالجة الحاسوب للأفعال داخل سياق معين.

4. إشكالات المعالجة الآلية للأفعال الخاصة في العربية

نتبنى في معالجة الأفعال الخاصة "كان وأخواتها" آليا ما قدمه النحاة بشأنها، إذ عملوا على تقسيمها إلى ثلاثة: يتصرّف القسم الأول تصرّفاً تاماً، (يجيء منه الماضي، والمضارع، والأمر)، وهو سبعة أفعال: (كان- أصبح- أمس-أضحى -بات -صار- ظلّ). ويتصرف القسم الثاني تصرفاً ناقصاً (يجيء منه الماضي، والمضارع فقط) ، وهو أربعة أفعال: (مازال-ما برح -ما فتئ – ما انفكّ)؛ فلا يأتي منها الأمر، لأن من شروط عملها أن يسبقها نفي، وهو لا يدخل على الأمر ولا يأتي منها المصدر لعدم دلالتها على الحدث. أما القسم الأخير فهو جامد (لا يتصرف مطلق)، ملازم لصورة الماضي، وهما فعلان) ليس– مادام (العزيز، 2019، صفحة 319).

1.2. مراحل التحليل الحاسوبي

لتجاوز الإشكال، المذكور آنفا، الذي يعترض معالجة بعض ظواهر اللسان العربي آليا، عملنا على إدراج الأفعال الخاصة

[5] نقصد بكل نوع من ناحية اللزوم والتعدية؛ أي أن الفعل اللازم يكتفي فقط بالشق الاساسي لتمام معناه أي الفاعل فقط، أما النوع الثاني (الفعل المتعدي) فبدوره يحتاج إلى الشق الثاني ليتم معناه رغم أنه حقق شرط ثنائية المسند والمسند إليه، إلا أن الشق الأساسي لابد له من شق ثانوي لإتمام معناه من الناحية الدلالية.

[6] هناك بعض الأفعال في اللسان العربي تتصرف تصرفا ناقصا؛ بمعنى أنها تتصرف فقط في المضارع والأمر مثل فعل" يَذَرُ" و"ذر"، ولا يجيء منه صيغة الماضي نحو قوله تعالى: - ﴿وَيَذَرَكَ وَءَالِهَتَكَ﴾[الأعراف:70]. / - ﴿اتَّقُوا اللَّه وَذَرُوا مَا بَقِيَ مِنَ الرِّبَا﴾ [الأنعام:12].

[7] - فـخ: فعل خاص - مب: مبتدأ - خب: خبر

[8] - فـت: فعل تام متعدي - فا: فاعل - مف: مفعول به

بجميع متغيراتها الصرفية واللواصق الممكنة التي يقبلها الفعل في منصة SAFAR platform[9]، وذلك، بتوليف الكلمات بناء على:

- مُجدع سفر _ SAFAR stemmer: وهو مورد اللواصق الممكنة لتوليف الكلمة العربية، بحيث يمكنه التعرف على الكلمة آليا وتحديد اللواصق التي تقبله.
- مورد المتصلات _ clitics resource: هو مورد للتعرف الآلي على التعريف والتنكير في اللسان العرب
- تقنية الترميز المعجمي _ lexical markup framework: وهي تقنية لتوفير المعطيات اللسانية لعمليات توليف الكلمات اللواصق.
- أدوات المعالجة الآلية للألسن الطبيعية: NLTK[10] وstandford NLP وغيرها من الأدوات التي توفر معطيات جاهزة لتوليفات الكلمة العربية مع اللواصق الممكنة.

وقد اعتمدنا في ذلك مجموعة من المراحل:

تتمثل **المرحلة الأولى** في إدخال الفعل الخاص "كان وأخواتها" معلوماتيا مع جميع اللواصق الممكنة في العربية التي تدخل على الفعل العربي؛ وهي خمسة سوابق، [همزة الاستفهام (أَ) وسين التسويف (سَ) وحرفا العطف الفاء والواو (فَ، وَ) واللام؛ لام التعليل(لِ)، لام القسم (لَ)]، أسفرت هذه المرحلة على توليد ثلاث وعشرين توليفة. أما إدراج اللواحق المتمثلة في ضمائر متصلة نحو كاف الخطاب وفروعها (كَ، كِ، كُمْ، كُمَا، كُنَّ)، وهاء الغائب وفروعها (هُ، هَا، هُمْ، هُمَا، هُنَّ) ونون الوقاية (نِي)، فولَّد ثلاث وأربعين توليفة انطلاقا من مبدأ الممكن وغير الممكن. وبناء على نظرية الاحتمالات وما هو مخزن بذاكرة الحاسوب، فإن عملية التوليف بين اللواصق والفعل الخاص لم تخضع لقواعد لسانية مما أدى إلى توليد التوليفات الصحيحة واللاحنة لعدم توفره على قواعد تسمح بتوليد السليمة منها فقط.

أما **المرحلة الثانية** فتتجلى في حصر كل التوليفات الممكنة في اللسان العربي، والتي بلغ عددها (380271) توليفة، من خلال ضبط الأفعال الخاصة حسب طبيعتها التصريفية إلى أقسام ثلاثة،[11]وتدقيقها لغويا. ثم فرز الصحيحة منها، والتي بلغ عددها (124479) توليفة، لتظل التوليفات المتبقية البالغة (25792 2) لاحنة، [12] كما هو موضح في الجداول أسفله:

[9] Software Architecture For Arabic language pRocessing
[10] Natural Language Toolkit
[11] - يتضمن القسم الأول: كان، أصبح، أضحى، أمسى، بات، صار، ظلّ.
- يتضمن القسم الثاني: ما انفك، ما برح، ما زال، ما فتئ.
- يتضمن القسم الثالث: لَيْس، ما دام.
[12]تستعمل لغة البرمجة لغة خاصة لفهم الحاسوب والتواصل معه، وللتمييز بين اللاحن والصحيح، وظفنا ثنائية من لغة البرمجة، فالرقم 0 يدل على اللاحن والرقم 1 يدل على الصحيح.

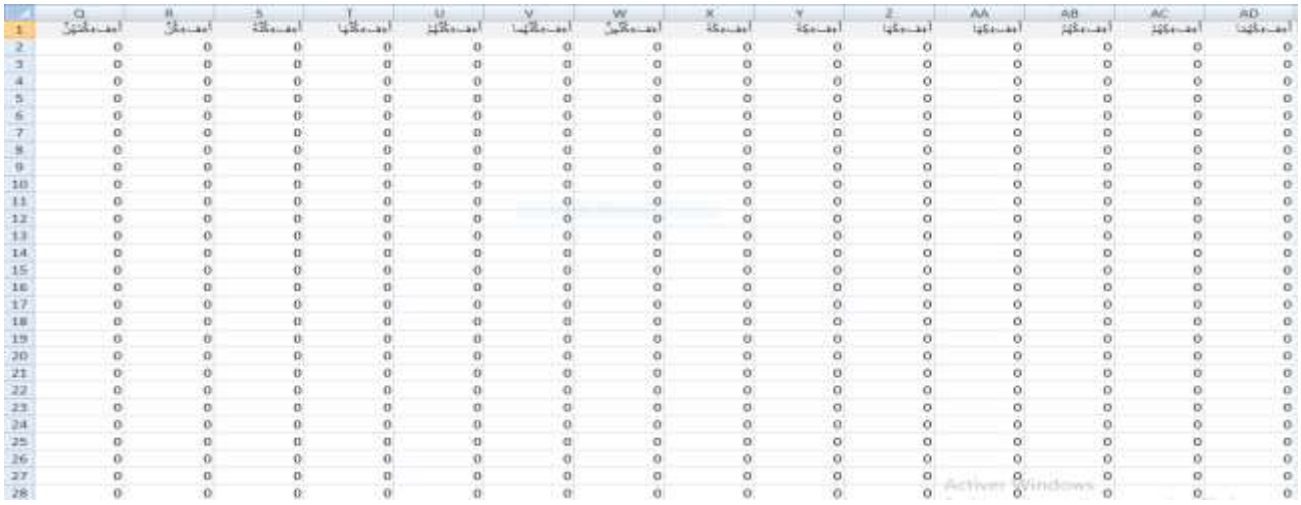
الجدول (1): توليفات القسم الأول من الأفعال الخاصة

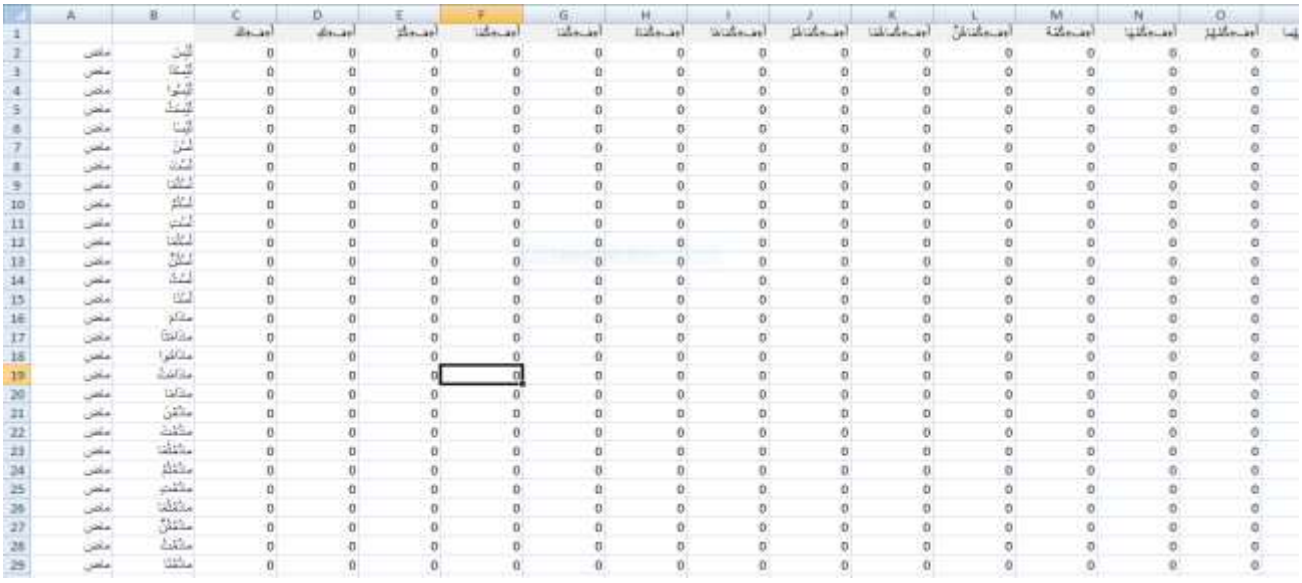
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1			أف+ك	أف+ي	أف+م	أف+نا	أف+نا	أف+نا	أف+نا+ا	أف+نا+هم	أف+نا+نا	أف+نا+هن	أف+نا	أف+نا	ن+هم
2		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
3		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
4		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
5		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
6		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
7		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
8		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
9		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
10		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
11		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
12		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
13		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
14		ماض	0	0	0	0	0	0	0	0	0	0	0	0	
15		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
16		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
17		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
18		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
19		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
20		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
21		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
22		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
23		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
24		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
25		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
26		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
27		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
28		مضارع	0	0	0	0	0	0	0	0	0	0	0	0	
29		امر	0	0	0	0	0	0	0	0	0	0	0	0	
30		امر	0	0	0	0	0	0	0	0	0	0	0	0	
31		امر	0	0	0	0	0	0	0	0	0	0	0	0	
32		امر	0	0	0	0	0	0	0	0	0	0	0	0	
33		امر	0	0	0	0	0	0	0	0	0	0	0	0	
34		امر	0	0	0	0	0	0	0	0	0	0	0	0	
35		ماض	0	0	0	0	0	0	0	0	0	0	0	0	

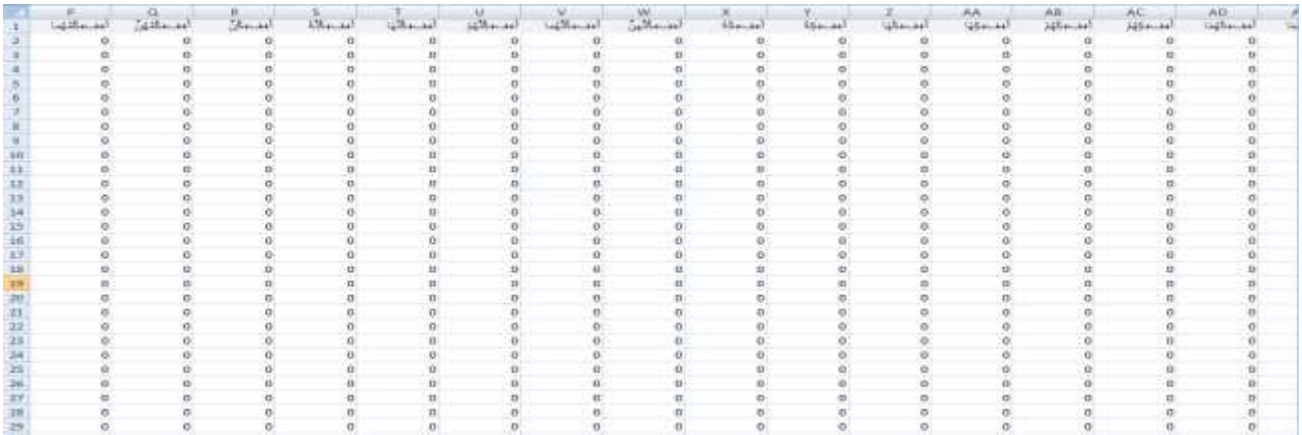
الجدول (2): توليفات القسم الثاني من الأفعال الخاصة

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1			أف+ك	أف+ي	أف+م	أف+نا	أف+نا	أف+نا	أف+نا+ا	أف+نا+هم	أف+نا+نا	أف+نا+هن	أف+نا	أف+نا
2		ماض	0	0	0	0	0	0	0	0	0	0	0	0
3		ماض	0	0	0	0	0	0	0	0	0	0	0	0
4		ماض	0	0	0	0	0	0	0	0	0	0	0	0
5		ماض	0	0	0	0	0	0	0	0	0	0	0	0
6		ماض	0	0	0	0	0	0	0	0	0	0	0	0
7		ماض	0	0	0	0	0	0	0	0	0	0	0	0
8		ماض	0	0	0	0	0	0	0	0	0	0	0	0
9		ماض	0	0	0	0	0	0	0	0	0	0	0	0
10		ماض	0	0	0	0	0	0	0	0	0	0	0	0
11		ماض	0	0	0	0	0	0	0	0	0	0	0	0
12		ماض	0	0	0	0	0	0	0	0	0	0	0	0
13		ماض	0	0	0	0	0	0	0	0	0	0	0	0
14		ماض	0	0	0	0	0	0	0	0	0	0	0	0
15		ماض	0	0	0	0	0	0	0	0	0	0	0	0
16		مضارع	0	0	0	0	0	0	0	0	0	0	0	0
17		مضارع	0	0	0	0	0	0	0	0	0	0	0	0
18		مضارع	0	0	0	0	0	0	0	0	0	0	0	0
19		مضارع	0	0	0	0	0	0	0	0	0	0	0	0
20		مضارع	0	0	0	0	0	0	0	0	0	0	0	0
21		مضارع	0	0	0	0	0	0	0	0	0	0	0	0
22		مضارع	0	0	0	0	0	0	0	0	0	0	0	0
23		مضارع	0	0	0	0	0	0	0	0	0	0	0	0
24		مضارع	0	0	0	0	0	0	0	0	0	0	0	0
25		مضارع	0	0	0	0	0	0	0	0	0	0	0	0
26		مضارع	0	0	0	0	0	0	0	0	0	0	0	0
27		مضارع	0	0	0	0	0	0	0	0	0	0	0	0
28		مضارع	0	0	0	0	0	0	0	0	0	0	0	0

الجدول (3): توليفات القسم الثالث من الأفعال الخاصة







تُبيِّن المعطيات أعلاه نتائج عملية دمج وتوليف السوابق واللواحق مع كان وأخواتها بوصفها أنموذجًا للفعل الخاص. وقد استعملنا ثنائية (0-1)، فالرقم (**0**) يدل على اللاحنة منها والرقم (**1**) يدل على الصحيحة؛ حيث تم إدراج كل اللواصق العربية الممكنة مع الفعل الخاص مصرف إلى الماضي والمضارع والأمر، فأنتج الحاسوب كل التوليفات الممكنة وغير الممكنة استنادًا على نظرية الاحتمالات وبناء على المعطيات المتوفرة في بياناته عن الفعل التام المتعدي، فعلى سبيل المثال قدمنا الفعل (كُنْتُ) والسابقتين (أَ، فَ) واللاحقة (كَ)؛ فأنتج لنا الحاسوب التوليفة (أَفَكُنْتُكَ*) وهي توليفة لاحنة.

في هذه الورقة حددنا كعينة لدراسة الأفعال الخاصة (كان وأخواتها) وقسمناها إلى ثلاثة أقسام؛ يتضمن القسم الأول: كان، أصبح، أضحى، أمسى، بات، صار، ظلَّ، ويتضمن القسم الثاني: ما انفك، ما برح، ما زال، ما فتئ، أما القسم الثالث فيتضمن: لَيْس، ما دام. وأثناء دمج السوابق مع (كان وأخواتها) دون الاعتماد على قواعد تضبط هذه التوليفات، أنتجت عملية الدمج مجموعة من الأخطاء نحو: لام التعليل (لِ) تدخل على الفعل المضارع فتنصبه، لكن توليفات الحاسوب رفعته (لِتكونُ*) وهي لاحنة والصحيح هو) لِتكونَ). أما أفعال القسم الثاني و"مادام " من القسم الثالث فلا تقبل لام التعليل وسين التسويف بسبب "ما النفي والمصدرية" التي أُضيفت إليه، كما لا تقبل أيضا كُلًّا من سين التسويف ولام التعليل ولام القسم، مثل: (سَمَازالَ*)_(لِمازالَ*)_ (لَمايَزولَنَّ*).

نشير كذلك إلى مثال آخر وهو أحد أخوات كان (ليس)، فإذا قدمنا له سين التسويف (سَ) وهمزة الاستفهام (أَ) فإنّ هذا يعطينا التوليفة: (أَسَلَيْسَ*)؛ اعتبرها الحاسوب توليفة ممكنة، إلا أنها غير ممكنة دلاليا في اللسان العربي لأن سين التسويف تدخل على الفعل المضارع، في حين أن الفعل (لَيْسَ) فعل جامد لا يتصرف إلى المضارع وبالتالي لا يقبل السوابق التي تدخل على الفعل المضارع فقط كسين التسويف.

نخلص إلى نتيجة مفادها أن الفعل الخاص لا يقبل جميع اللواصق وخاصة اللواحق بسبب فراغه من سمة الحدث في نفسه ولا يؤدي وظيفته إلا في سياق المقولات المعجمية المكونة من مبتدأ وخبر. وإذا ما قورنت هذه النتيجة بفعل تام متعدي مثل: (أَفَأَخذتُكَ)، فهذا الأخير صحيح من ناحية التوليف والمعنى لأن بنيته تتكون من زمن الماضي وحدث الأخذ؛ لذلك أنتج لنا توليفة صحيحة، عكس الفعل الخاص الذي لا يقبل اللواحق كما هو مبين في المعطيات أعلاه. وبناء على المعطيات المذكورة أعلاه، فالحاسوب يواجه إشكالات عديدة أهمها أن:

- الفعل المتعدي يحتاج مركبين اسميين (الفاعل والمفعول به) لإتمام المعنى، كما يحتاج الفعل الخاص أيضا لمركبين اسميين اسم وخبر)لإتمام المعنى، فكيف للآلة أن تميز بين الأفعال التامة وخاصة المتعدية منها والأفعال الخاصة في سياق الجملة؟
- ما السبب الحقيقي في الاختلاف بين الفعل التام المتعدي والخاص؟ وكيف للحاسوب أن يفقه هذا الاختلاف حتى لا يقع في اللُّبس مما يجعله ينتج جملا لاحنة؟

أما **المرحلة الثالثة** التي خصصناها لطرح الإشكالات التي تحول دون تعرف الحاسوب على الأفعال الخاصة، ورصدنا كذلك مكامن الخلل بغية تفسيرها افتراضا وقاعدة واستنتاجا، والتي تتجلى في أن:

- الفعل الخاص لا يتضمن حدثًا كسائر الأفعال التامة، لذلك لا يُمكِنُ أن يرِد مبنيا للمجهول ولا يصاغ منه اسم الفعل أو اسم الفاعل أو اسم المفعول وغيرها من المشتقات التي تُشتَق من الفعل التام المتعدي بسلاسة.
- سبب لحن التوليفات يرجع لعدم وجود سمة الحدث بالفعل الخاص وهو ما يجعل دمج اللواحق بالفعل لاحنة من ناحية المعنى.
- إضافة سمة [±حدث] هو حلٌّ جذري لتجاوز إشكال التمييز بين الفعل الخاص والفعل التام المتعدي آليا. كما ستسمح إضافة هذه السمة من التعرف على الفعل الخاص في سياق معين.
- الحاسوب لا يميز بين الفعل التام المتعدي والفعل الخاص. لذلك، فدمج اللواصق مع الفعل الخاص، ينتج توليفات بناء على خصائص الفعل التام المتعدي الموسومة بسمة {[+زمن]، [+حدث]}، بينما يوسم الفعل الخاص (كان وأخواتها) بسمة {[+زمن]، [-حدث]}.
- غياب سمة الحدث من الفعل الخاص هو السبب الرئيسي في عدم تقبله اللواحق أثناء دمجه مع اللواصق من ناحية المعنى.

5.خاتمة

يتميز اللسان العربي بخاصية الاشتقاق، فجذر واحد من ثلاثة حروف أصول يشتق منه عدد لا حصر له من الكلمات. فالأصل في الاشتقاق هو الفعل التام الذي يتوفر على سمتي حدث وزمن، وتشتق منه الفاعل والمفعول واسم الفاعل واسم المفعول وغيرها، إلا أن هذا الشرط لا يتوفر في الأفعال الخاصة (كان وأخواتها وكاد وأخواتها). ومن ثم فغياب سمة الحدث يؤثر في العديد من الجوانب أهمها الجانب الاشتقاقي وعملية الدمج.

تمثل المعالجة الآلية للأفعال الخاصة "كان وأخواتها" مرحلة متقدمة لأنها تستلزم وضع قواعد مبرمجة حاسوبيا من أجل تجاوز كل الإشكالات التي تم رصدها انطلاقا من نتائج البحث المستمر[13] لدراسة الأفعال الخاصة في اللسان العربي. كما اقتضت هذه الدراسة معالجة عدد هائل من التوليفات انطلاقا من ست قواعد، فقد عالجت القاعدة الأولى إشكالية التعرف الآلي على الفعل الخاص مقارنة بالفعل التام المتعدي، وذلك بإضافة سمة مميزة للفعل الخاص. في حين عالجت القاعدة الثانية منها (251093) توليفة لاحنة، تضمنت الفعل الخاص مع اللواحق، وهو ما لا يقبله الفعل الخاص. أما الثالثة، فعالجت (497) توليفة، تضمنت حركة الرفع الإعرابية في السابقة لام التعليل، لكن هذه الأخيرة تعمل بالنصب. وعالجت القاعدة الرابعة (924) توليفة تضمنت الفعل الخاص بالقسم الثاني المبتدئ بما النفي، وتبين لنا هذه القاعدة لحن سين التسويف عند توليفها مع الأفعال الخاصة نحو: (سَمَايَزَالُ*). وفيما يخص القاعدة الخامسة، فقد عالجت (792) توليفة لاحنة؛ تضمنت حركة النصب الإعرابية لـلام التعليل. وتوضح هذه القاعدة عدم إمكانية دخول لام التعليل على القسم الثاني من الأفعال الخاصة. وأخيرا عالجت القاعدة السادسة (2486) توليفة، تضمنت لام القسم وتأثيرها باللاحقة التي تضاف للفعل الخاص بالمضارع، وهي نون التوكيد اللازمة للقسم.

لائحة المصادر والمراجع

المصادر والمراجع العربية

- الأندلسي، جمال الدّين2001م. *شرح التسهيل: تسهيل الفوائد وتكميل المقاصد*. تحقيق عطا، محمد عبد القادر-السيّد طارق فتحي، بيروت-لبنان: دار الكتب العلمية، الطبعة الأولى.
- الدّحداح، أنطوان 1991م. *معجم تصريف الأفعال العربية*. راجعه د. جورج متري عبد المسيح، إعادة الطبع 1995، بيروت_ لبنان: مكتبة لبنان ساحة رياض الصلح، الطبعة الأولى.
- الدخيسي، عبد الكريم 2020م. *اللسانيات الحاسوبية والمعالجة الآلية للغة العربية*. تقديم إدريس مقبول، عالم الكتب الحديث للنشر والتوزيع إربد، الأردن، الطبعة الأولى.
- رشوان، محسن وآخرون 2019م. *مقدمة في حوسبة اللغة العربية*. الرياض_ المملكة العربية السعودية: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، الطبعة الأولى.
- شني فليح خضير2014. *مفهوم الكلمة بين اللغويين القدماء والمحدثين*. مجلة لارك للفلسفة واللسانيات والعلوم الإجتماعية، العدد الثالث عشر، السنة السادسة، جامعة واسط (كلية الآداب)، الكويت، العراق.
- عبد العزيز، خالد 2019، *النحو التطبيقي*. المنصورة-مصر: دار اللؤلؤة للنشر والتوزيع، الطبعة الثالثة.
- الغامدي، منصور بن محمد وآخرون 2017م. *مدخل إلى اللسانيات الحاسوبية*. الرياض_ المملكة العربية السعودية: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، الطبعة الأولى.
- الفاسي الفهري، عبد القادر1990م. *البناء الموازي: نظرية في بناء الكلمة وبناء الجملة*. الدار البيضاء-المغرب: دار توبقال للنشر، الطبعة الأولى.
- الفاسي الفهري، عبد القادر 2007. بمشاركة، العمري، نادية: *معجم المصطلحات اللسانية: إنجليزي-فرنسي-عربي*. الرباط: دار الكتاب الجديد المتحدة.
- قنبر، أبي بِشر عمرو 1408ه- 1988م. ***كتاب سيبويه***. تحقيق عبد السلام محمد هارون، القاهرة: مكتبة خانجي، الطبعة الثالثة.
- مبارك، مبارك 1995. *معجم المصطلحات الألسنية: فرنسي-إنكليزي- عربي*. بيروت: دار الفكر اللبناني، الطبعة الأولى.

المصادر والمراجع باللغة الأجنبية

- Jack C. Richards & Richard Schmidt. 2010. *Dictionary of language teaching & applied linguistics*, 4th edition, Rouledge.
- Alajmi & others. *Towards an arabic stop-words list generation*, International Journal of Computer Applications, Volume 46– No.8, 2012, Helwan university, (2012).
- Nazief, B. *Development of Computational Linguistics Research: a Challenge for Indonesia*. Computer Science Center. University of Indonesia Jakata, Indonesia. (https://www.researchgate.net> publication, 21/03/2020).
- Sadghi, M & Vegas, J. *Automated identification of light stop words for Persian information retrieval systems*, journal of information science. computer science department: university of Valladolid, (Published by: https: www. jis.sagepub.com, (2014).
- Tsujii, J. *Computational linguistics and Natural Language Processing*. Conference Paper. February 2011. Pp 268-282. Published by: (https://www.researchgate.net/publication/2270737

[13] ونشير إلى أننا مازلنا في طور البحث والدراسة الكلمة الخاصة في سلك الدكتوراه ولا زلنا نشتغل على نفس الموضوع.