

مقاربة صوتية للتوليد الآلي للجذور الثلاثية في اللسان العربي المعياري
**A Phonological Approach to Automatic Generation of
 Trilateral Roots in Standard Arabic Language**

المودن محمد*

جامعة محمد الخامس بالرباط، كلية الآداب والعلوم الإنسانية (المغرب)

mohamed_elmoudden@um5.ac.ma

تاريخ النشر: 2023/12/31	تاريخ القبول: 2023/12/08	تاريخ الإرسال: 2023/10/06
-------------------------	--------------------------	---------------------------

الملخص:

يتأطر هذا المقال ضمن الجهود اللسانية الرامية إلى فهم آليات تشكيل المفردة العربية، انطلاقاً من مبادئ الصوتية اللاخطية. ونهدف من خلاله إلى بناء برنامج آلي مولد للجذور الثلاثية في اللسان العربي المعياري، وهو برنامج يحاكي ما يحدث في دماغ المتكلم العربي، من منظور اللسانيات التوليدية؛ لأنه لا يعتمد على جذور مخزنة في الدماغ على شكل معاجم، بل يشتغل بدالة رياضية بسيطة. تمكن من التمييز بين الجذور الثلاثية التي يسمح بها نظام اللسان العربي المعياري، والجذور التي لا يسمح بها هذا النظام. إذ لا يخفى على الباحثين أننا ما زلنا نجهد كيفية بناء المفردة العربية، على المستوى الصوتي.

كلمات مفتاحية: توليد آلي؛ جذر ثلاثي؛ صوتية لاخطية؛ خوارزميات؛ لسانيات.

Abstract:

This article is a component of linguistic research targeted at comprehending the mechanics of Arabic word production based on non-linear phonological principles. In order to model what occurs in an Arabic speaker's brain from a generative linguistics perspective, we want to develop via it an automated program for

generating trilateral roots in standard Arabic. It uses a straightforward mathematical formula to function rather than the linguistic roots that are kept in dictionaries in the brain. It has the ability to discriminate between trilateral roots that are authorized by the Arabic linguistic system and those that are not.

As is well known, the phonological makeup of Arabic words continues to be a source of discussion among experts. This is a serious issue that necessitates more study as well as the suggestion of novel processes.

Keywords :Automated generation; trilateral root; non-linear phonology; Algorithms; Linguistics.

1. مقدمة:

يجد المهتم بمشاريع حوسبة الكلمة العربية أن صعوبات تتعلق بفهم بنية الكلمة قد اعترضتها. الأمر الذي انعكس سلبا على مشاريع حوسبة المفردة العربية؛ حيث تعتمد جل البرامج التي اطلعنا عليها إلى تخزين المعاجم العربية الضخمة في قواعد بيانات مدرجة مع البرنامج، ثم تقوم باستدعاء الكلمة أو ما يتعلق بها. وهكذا يصعب أن نتحدث عن إمكانية توليد المفردة في اللسان العربي آليا، أو على الأقل نادرة هي البرامج التي اتجهت إلى البحث في إمكانية توليد المفردات العربية انطلاقا من صورة قواعد الصرف العربي القديم على شكل خوارزميات، ذات دَخلٍ وخَرجٍ، غير أن هذا النحو من الأبحاث تواجهه إشكالات مرتبطة بقضايا صرفية عالقة في الدرس اللغوي العربي القديم، مثل المصادر الثلاثية المجردة، وجموع التكسير، وغيرها من المفردات المحمولة على السماع، ولم يُتوصل بعدُ إلى القواعد الضابطة لتكوينها.

إن تجاوز ما سلف يتطلب، في نظرنا، تقدّم البحث الصرف-صواتي وتطويره لفهم بنية الكلمة العربية حتى يتسنى لنا فتح آفاق جديدة لمشاريع الحوسبة في اللسانيات العربية، وتبَيّ مشاريع تنسجم وطبيعة اللغة في الدماغ البشري، إذ نعتقد، انطلاقا من فرضيات اللسانيات التوليدية، أن اللغة لا تُخزن في دماغ المتكلم، على النحو الذي تُخزّن فيه المعاجم في قواعد

البيانات، وإنما يخزن الدماغ نسق اللغات الطبيعية، ثم يولد ما يسمح به هذا النسق، من خلال قواعد صورية محددة؛ وقد أظهرنا في المودن (2022) أن المدخل الصرف-صواتي التوليدي من المداخل اللسانية التي يمكن أن تدفع بالبحث الصرفي العربي نحو إمكانية استنباط قواعد جديدة للصرف العربي، ومنها استنباط الخوارزميات المشابهة للنسق العربي الموجود في ذهن المتكلم بالعربية. وهذا ما سيتطلب تطوير نظريات صوتية على النحو الذي نجده في التاقي (2021)، لبلورة نظرية شاملة تجيب عن العلاقات القائمة بين القطع الصوتية المشكلة للجذر، قبل الانتقال إلى التفكير في العلاقات القائمة بين الجذر والصيغ الصرفية.

وبناء على ما سلف يأتي هذا المقال كمحاولة لبناء برنامج يحاكي ما يحدث في دماغ المتكلم العربي ليقوم بتوليد الجذور الثلاثية التي يسمح بها نسق اللسان العربي المعياري، بناء على قواعد صورية، على شكل دوال رياضية، مستنبطة من دراسة صرف-صواتية لأليات تشكيل المفردة العربية. ومن هنا تتأتى أهميته، وسنعمل على تحويل استنتاجات القواعد الصوتية التي تحكم تكوين الجذر إلى خوارزميات للتوليد الآلي للجذور الثلاثية في اللسان العربي المعياري. ثم ترجمة هذه الخوارزميات إلى أكواد برمجية بلغة بايثون لبلورة برنامج آلي مولد لهذا النمط من الجذور.

يقتضي منا كل ما سلف أن ينتظم هذا المقال وفق مقدمة، ثم مبحثين، نعالج في أولهما خوارزميات التوليد الآلي للجذور الثلاثية في اللسان العربي المعياري، ثم توصيف للمولد الآلي لهذه الجذور في مبحث ثان، وفي كل ذلك نتوسل بمناهج متعدد تقتضيها طبيعة ما نعالجه في هذا المقال؛ إذ سنكون بحاجة إلى المنهجين الوصفي والتفسيري، بالإضافة المنهج النقدي.

2. خوارزميات التوليد الآلي للجذور الثلاثية في اللسان العربي

1.2 في نقد خوارزميات المعالجة الآلية للسان العربي المعياري:

كشفنا في المودن (2022) أن مشاريع حوسبة الكلمة العربية اعترتها صعوبات متعلقة بفهم بنيتها، وهو ما جعل جملها يعتمد على قواعد بيانات تخزن فيها الكلمات العربية، ثم تعتمد على الاستدعاء لا التوليد، دون أن ننكر أن مشاريع أخرى اتجهت إلى توليد المفردات العربية

انطلاقاً من صورنة قواعد الصرف العربي القديم على شكل خوارزميات، ذات دَخلٍ وخَرجٍ، لكنها تواجه إشكالات مرتبطة بقضايا صرفية، حملت في النمذجة اللغوية العربية القديمة على السماع.

وحتى نضع القارئ الكريم أمام واقع المقاربات المختلفة في حوسبة الصرف العربي خاصة، نقدم نظرة موجزة عن أهم مشاريع حوسبة الكلمة العربية التي نذكر منها مشروع برنامج الخليل الصرفي الذي يسعى إلى تحديد الصورة الصرفية للكلمة، وذلك بالتعرف على بنيتها، ومن ثمة استدعاء المفردات التي تشترك في الجذر ذاته، وهذا المسعى يقتضي تحليل الكلمة حين يستقبلها سواء أكانت مشكولة كلياً، أم جزئياً أم غير مشكولة، ثم يستقبلها مفردة أو داخل التركيب، بحيث يتعامل مع الكلمات المشكلة للجمل أو النص على أنها كلمات مفردة مستقلة، ويعالجها تبعاً؛ أي أنه يعالج الكلمة خارج النص والسياق، مما يجعله ينظر إلى الكلمة من أوجه مختلفة، لكون الكلمات العربية حمالة أوجه، ما لم تشكل أو تكون داخل سياق يحدد معناها¹.

ونجد كذلك مشروع معالج صرفي متعدد الأطوار (برنامج صخر للتحليل الصرفي) الذي يهتم بالكلمة المفردة؛ حيث صمم ليغطي نطاق الكلمات العربية بالكامل؛ ويعتمد الجذر مدخلاً للتعرف على المفردة العربية؛ إذ يقوم بالتوصل إلى جذر الكلمة بعد تجريدها من كافة اللواحق واللواحق. ثم يقدم بيانات صرفية للمفردة من قبيل: الجذر، والوزن... وهي العناصر نفسها التي يعيد من خلالها توليد الكلمات العربية بالإضافة إلى التأكد من صحة الكلمة بإحداث مقارنة بين الناتج التوليدي والكلمة المخزنة لديه في قاعدة المعطيات².

ونختتم هذه المشاريع بمشروع المحلل الصرفي لتيم بكوالتر الذي يعدّ من أبرز البرامج المشهورة في مجال المعالجة الآلية للكلمة العربية، وقد اعتمد عليه اتحاد البيانات اللغوية في تصنيف الكلمات في النصوص العربية. ويقوم على إعادة كتابة الحروف العربية بحروف لاتينية. وهو على خلاف باقي البرامج؛ إذ يعتمد على الجذع بدل الجذر. ويكون بذلك يحاكي برامج معالجة الألسن ذات الطابع الجذعي (الصرف السلسلي) كالإنجليزية والألمانية³

إن المشاريع السابقة، رغم أهميتها، تتطلب تقدّم البحث الصوتي في فهم بنية الكلمة العربية حتى يتسنى لنا فتح آفاق جديدة لمشاريع الحوسبة في اللسانيات العربية، وتبني مشاريع تنسجم وطبيعة اللغة في الدماغ البشري، كما تصور ذلك اللسانيات التوليدية. وإن هذا المسعى لن يتحقق، في نظرنا، ما دمنا نحاول معالجة اللسان العربي، بناء على ما قدمه الدرس اللغوي العربي القديم، ليس لنقص فيه بل لاختلاف الوسائل والغايات بين الأمس واليوم.

وقد نتجت عن ذلك إشكالات تحول دون تحقيق معالجة آلية فعالة للسان العربي المعياري؛ الأمر الذي دفع المهتمين بهذا المجال إلى التعبير عن القصور الذي شاب البحث الآلي في اللسان العربي حديثاً. كما عبر عن ذلك المهيوبي (2018)؛ حين تطرق لهذه الإشكالات التي نجمل أبرزها في الآتي:⁴

- ضعف الدعم المقدم للشركات، ومراكز البحث العلمي لتطوير برامج لمعالجة العربية حاسوبياً.
- بعثرة الجهود العربية، سواء على المستوى النظري أم التطبيقي: فكل باحث وكل منظمة تعمل بمعزل عن غيرها. إضافة إلى محدودية الدراسات في هذا المجال، فإنها تعاني من ضعف الانتشار، وانعدام التكامل والتعاون بينها، ولا تكاد تتجاوز الملتقيات والندوات.
- افتقار معظم الحاسوبيين العاملين في حوسبة العربية إلى الحد الأدنى من المعرفة اللغوية، لذلك ينبغي أن تتضافر جهود اللغويين والحاسوبيين في أي مشروع علمي يهدف إلى معالجة اللغة العربية آلياً.
- الطابع التجاري الذي أصبح يحكم ضرورة الإنجاز الحاسوبي، وجعل من حقل اللسانيات الحاسوبية حقلاً تجارياً يخضع لسوق العرض والطلب.

إن من شأن هذا الواقع أن يدفع الباحثين في اللسان العربي المعياري إلى استيعاب أن البحث اللغوي العربي عليه أن ينظر إلى اللغة على أنها "ليست نظاماً عشوائياً، بل منظومة متسقة تقيدها الضوابط وتحكمها القواعد المطردة"⁵، ونعتقد أن ذلك يستوجب الانفتاح على الدرس اللساني المعاصر، ومحاولة تكييفه مع معطيات اللسان العربي، أو على الأقل الاشتغال

بمناهجه التي تقر بأن اللغة أشبه ما تكون، من حيث بنيتها العميقة، بالحاسوب. ورفض وجود قضايا صرفية لا تخضع للقياس⁶، وهذا الفكر نستطيع أن ندفع بدراسة العربية لتساير تحديات القرن الحادي والعشرين⁷.

وتحقيقاً لهذه الغايات نقترح أن تكون اللسانيات التوليدية مع تشومسكي مدخلا للتفكير في حوسبة المفردة العربية؛ حيث تقر اللسانيات التوليدية على أن المتكلم حين يروم إنتاج الكلام لا يستدعي الكلمات من معجم مخزن في ذاكرته، بل ثمة ضوابط تملها البنية العميقة على الجهاز النطقي، وهو ما فسره تشومسكي بكون المتكلم لا يعيد تكرار ما سمعه أثناء اكتسابه للغة، إنما يتميز بالقدرة على التوليد، أي توليد عدد غير متناه من الجمل بناء على قواعد متضمنة في الملكة اللغوية.

ومن هذا المنطلق يجب أن نبحت في هذه القواعد التي نراها عبارة عن خوارزميات، وليست القواعد التي يتعلمها الطلاب في المدارس؛ بحيث إن للإنسان قدرة على اكتساب لغة ما اكتساباً سليماً دون أن يدرس القواعد، فأثناء سماعه لأي لسان طبيعي (عربي، انجليزي...) يتعرف ذهنه على نسق هذا اللسان فيكتسبه. ويرى تشومسكي أن جزءاً هاماً من هذا النسق تتشاركه كافة الألسن الطبيعية البشرية⁸.

ومن هذه اللسانيات التوليدية ننتقي الصوتات اللاخطية لتكون أساساً لتطوير البحث الصرفي العربي، ولذلك ما يبرره، صوتياً وحاسوبياً؛ إذ يعول على الصوتات التوليدية لتسهم في تطوير البحث الصرفي العربي انطلاقاً من تقديم نتائج قابلة للصورنة، وإفهامها للآلة. خاصة بعد الإقرار بأن "الدراسات الصرفية القديمة غير كافية لبناء محلل صرفي حاسوبي للغة العربية"⁹.

ويشتغل المقال، بشكل دقيق، ضمن الصوتات المستقلة القطع، داخل الصوتات اللاخطية، لينظر في القوانين الصوتية، خاصة مبدأ المحيط الإيجاري (ocp)¹⁰ الذي يمنع تجاوز عنصرين متماثلين على مستوى المقطع ذاته¹¹، وبالضبط على مستوى خط الجذر لأن المفردة في الصوتات اللاخطية بات يمثل لها على مستوى متعدد الخطوط بعد أن ظهرت

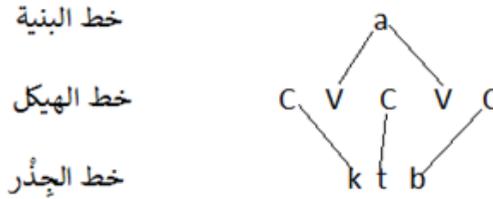
النظرية مستقلة القطع مع كولدسميث (1976-1990-2016) Goldsmith ومككارثي (1979)، (1981) McCarthy، وليبن (2017) Leben باعتبارها إحدى النظريات التي اقترحت تجاوزا للنظرية الصوتية المعياري¹²، مع تشومسكي وهالي (1968) Chomsky & Halle القائمة على التحليل الآلي للأسن الطبيعية اعتمادا على تمثيل صوتي وفونولوجي لأصوات الألسن الطبيعية من أجل بناء نظام من القواعد لتوليد عناصر صوتية انطلاقا من بُنى عميقة ومجردة.

إن مبدأ المحيط الإجباري، كقانون صوتي، يقودنا إلى ضبط العلاقات القائمة بين العناصر الصوتية (الصوامت/عناصر الجذر) لتشكيل جذر المفردة في اللسان العربي المعياري. وبعبارة أخرى إن تطبيق مبدأ المحيط الإجباري يجري على مستوى البنية العميقة (أي ما يحدث في ذهن المتكلم) ليعطينا على مستوى الخَرْج (التحقق الصوتي) كلمة مكونة من جذر صوتي يحكم تشكيله مبدأ صوتي. وقيمة هذا المقال تكمن في البحث في تحويل هذه العمليات إلى الخوارزميات التي من خلالها نستطيع بناء مولد آلي محاك لما يحدث في ذهن المتكلم العربي.

2.2 خوارزميات لتوليد آلي للجذور الثلاثية في اللسان العربي:

لا يتسع المقام لعرض كل الخطوات المتبعة لاستنباط القوانين الصوتية المتحكمة في تكوين الجذور الثلاثية في اللسان العربي المعياري، غير أنه خلصنا، سلفا، إلى أن مبدأ المحيط الإجباري هو المسؤول عن ضبط العلاقات القائمة بين عناصر الجذر في اللسان العربي المعياري. كما خلصنا إلى أننا في الصوتية مستقلة القطع أصبحنا نتحدث عن مستوى يقترح نموذجا لتحليل المفردة، يقتضي "وجوب تجزئ التمثيل المعياري الأحادي الطبقة إلى العديد من الطبقات. تشكل كل طبقة ترتيبا خطيا للقطع، وتُربط القطع ببعضها البعض من مختلف الطبقات"¹³، كما نوضح ذلك في التمثيل الآتي (1).

(1) التمثيل الصوتي للمفردة في الصوتية اللاخطية



نلاحظ أن التمثيل للكلمة /katab/ جرى على ثلاثة خطوط¹⁴؛ خط للهيكل، يتم ملؤه عن طريق الصوامت (الحروف الأصول) والصوائت (الحركات)، وهو عبارة عن صورة مجردة لمواقع نغمية أو قطعية. ثم خط للجذر يتكون من الصوامت التي تتكون منها الكلمة، وهي التي يصطلح عليها في الأدبيات اللغوية العربية القديمة بالحروف الأصول، وأخيراً خط البنية، وهو خط عبارة عن صائت (a) الذي يطلق عليه الفتحة، ونسجل أيضاً أن هذا الصائت استُدرِج فملاً العنصر الرابع في خط الهيكل.

وهكذا فإن المفردة في الصوتية المستقلة القطع تُبنى من خلال طبقات أو خطوط كما في (1)، يُمثّل لها في المعجم بواسطة عقود الربط التي تسعفنا في ربط مقاطع المفردة؛ بتطبيق قواعد الربط، بدءاً من اليسار إلى اليمين، وتحديد العناصر المؤهلة للربط، أو الانتشار. وأن كل عنصر لا يُربط على مستوى المقطع لا يتحقق على مستوى النطق¹⁵. ومن مميزات هذا التمثيل الصوتي أنه قابل للصورنة، وبالتالي إمكانية صياغته على شكل خوارزميات.

والحديث عن الخوارزميات هو حديث عن مجموعة من الخطوات المنطقية المحددة، التي تستخدم لحل مشكلة أو إنجاز مهمة معينة. وهي من أساسيات الحوسبة والبرمجة، وتستخدم بشكل واسع في العديد من المجالات التي تتطلب تحليل البيانات والمعلومات، مثل الذكاء الاصطناعي وتعلم الآلة ... كما تتميز الخوارزميات بأنها تتبع خطوات منطقية محددة، وبالتالي يمكن تطبيقها على مجموعة متنوعة من البيانات لحل إشكالات تبدو معقدة¹⁶.

وخلاصة القول تُعدّ الخوارزميات من الأدوات الأساسية في مجالات علوم الحاسوب والذكاء الاصطناعي وتعلم الآلة، وتوفر طرقاً فعالة وسريعة لحل المشاكل المعقدة والتعامل مع البيانات الكبيرة. وعلى هذا الأساس، يمكن تحويل الاستنتاجات المتعلقة بتطبيق مبدأ المحيط

الإجباري لفهم العلاقات القائمة بين العناصر المكونة للجذر في اللسان العربية إلى خوارزميات، ذات تمثيل صوري؛ وهي التي يمكن ترجمتها إلى أكواد برمجية سعياً لبناء برنامج مولد للجذور الثلاثية في اللسان العربي المعياري.

ولقد أدى بنا فحص الجذور الثلاثية للسان العربي كما وردت في معجم لسان العرب لابن منظور؛ انطلاقاً من استحضار مبدأ المحيط الإجباري (ocp) في جزئه العام والخاص، إلى الخلوص إلى قواعد نصوغها على شكل خوارزمية لغوية، بعد التفكير في خطوات إنشاء المولد الآلي للجذور الثلاثية التي يسمح بها نسق اللسان العربي المعياري، كالاتي:

- الخطوة الأولى: تعريف توليد الجذر الثلاثي؛ ونعني به الجمع بين ثلاثة حروف في ترتيب معين يشكل الجذر الثلاثي، وفق قواعد صورية محددة سلفاً.

- الخطوة الثانية: تحديد الشروط اللازمة؛ إذ يجب على عملية توليد الجذر الثلاثي احترام مبدأ "ocp" الذي يحظر ما يأتي:

- يحظر تكرار حرفين متطابقين في الموقع الأول والموقع الثاني للجذر الثلاثي (قاعدة المبدأ العام).
- يحظر الجمع بين الواو والياء (و، ي) في الموقع الأول والموقع الثاني للجذر الثلاثي.
- يحظر الجمع بين الغين والخاء (غ، خ) في الموقع الأول والموقع الثاني للجذر الثلاثي.
- يحظر الجمع بين حرفين من الحروف الآتية: (ق، ج، ك) في الموقع الأول والموقع الثاني للجذر الثلاثي.
- يحظر الجمع بين حرفين من الحروف الآتية: (ل، ر، ن) في الموقع الأول والموقع الثاني للجذر الثلاثي.
- يحظر الجمع بين حرفين من الحروف الآتية: (ت، د، ط) في الموقع الأول والموقع الثاني للجذر.
- يحظر الجمع بين حرفين من الحروف الآتية: (ش، ض، س، ز، ص) في الموقع الأول والموقع الثاني.

- يحظر الجمع بين حرفين من الحروف الآتية: (ث، ذ، ظ) في الموقع الأول والموقع الثاني للجذر الثلاثي.
- يحظر الجمع بين حرفين من الحروف الآتية: (ف، ب، م) في الموقع الأول والموقع الثاني للجذر الثلاثي.
- يحظر الجمع بين الهمزة والهاء (ء، ه) في الموقع الأول والموقع الثاني للجذر الثلاثي.
- يحظر الجمع بين الحاء والعين (ح، ع) في الموقع الأول والموقع الثاني للجذر الثلاثي.
- الخطوة الثالثة: تطبيق الخوارزمية. لتوليد الجذر الثلاثي يفترض من المولد الآلي ما يأتي:
- يطلب من المستخدم إدخال ثلاثة حروف بالعربية.
- يتحقق من الشروط اللازمة لتوليد الجذر الثلاثي في العربية.
- إذا كانت الشروط مستوفاة، يتم توليد الجذر الثلاثي بشكل صحيح، ويتم طبعه للمستخدم. وإلا فإن العملية تفسل، وتطبع رسالة للمستخدم تقول: هناك قوانين صوتية تمنع من تكوين هذا الجذر في اللسان العربي المعيار، ويرجى إدخال ثلاثة حروف أخرى.
- ويمكن أن نعبر بصورة رياضية عن الشروط اللازمة لتوليد الجذر الثلاثي في اللسان العربي المعيار، كما في الدالة الرياضية الآتية:

(2) دالة رياضية تعبر عن قوانين تكوين الجذر الثلاثي في اللسان العربي المعيار

$$f(x) = \neg(x[0] = x[1]) \vee$$

$$(x[0] \in \{ 'و', 'ي' \} \wedge x[1] \in \{ 'و', 'ي' \}) \vee$$

$$(x[0] \in \{ 'غ', 'خ' \} \wedge x[1] \in \{ 'غ', 'خ' \}) \vee$$

$$(x[0] \in \{ 'ق', 'ج', 'ك' \} \wedge x[1] \in \{ 'ق', 'ج', 'ك' \}) \vee$$

$$(x[0] \in \{ 'ل', 'ر', 'ن' \} \wedge x[1] \in \{ 'ل', 'ر', 'ن' \}) \vee$$

$$(x[0] \in \{ 'ت', 'د', 'ط' \} \wedge x[1] \in \{ 'ت', 'د', 'ط' \}) \vee$$

$$(x[0] \in \{ 'ش', 'ض', 'س', 'ز', 'ص' \} \wedge x[1] \in \{ 'ش', 'ض', 'س', 'ز', 'ص' \}) \vee$$

$$(x[0] \in \{ 'ث', 'ذ', 'ظ' \} \wedge x[1] \in \{ 'ث', 'ذ', 'ظ' \}) \vee$$

$$\forall (x[0] \in \{ 'م', 'ب', 'ف' \} \wedge x[1] \in \{ 'م', 'ب', 'ف' \})$$

$$\forall (x[0] \in \{ 'ه', 'ء' \} \wedge x[1] \in \{ 'ه', 'ء' \})$$

$$((x[0] \in \{ 'ع', 'ح' \} \wedge x[1] \in \{ 'ع', 'ح' \}))$$

وتستخدم الدالة الشروط المذكورة لتحديد ما إذا كانت الحروف الثلاثة التي يدخلها المستعمل يمكن أن تكون جذرا ثلاثيا للمفردات في اللسان العربي المعيار، أو أن نظامه لا يقبل ذلك النمط من الصوامت لتكوين الجذر الثلاثي، فإذا تضمن الإدخال خطأ واحدا فإن عملية توليد الجذر لن تكتمل، فبحسب قوانين الصوامة اللاخطية التي يوفرها مبدأ المحيط الإيجباري، لا يوجد جذر ثلاثي لا تُعبّر عنه الدالة (2)، وإن وُجد فمن النادر أو الدخيل أو المعرب. مما لا يسمح المجال بالتفصيل فيه.

وخلاصة القول يمكن أن نلاحظ أن هذه الدالة تتضمن مجموعات صامتية، وتتدخل الدالة لتفحص العنصر الصامتي الذي يوجد على مستوى الموقع الأول والثاني للجذر المُدخل من طرف مستخدم البرنامج، وتتم إجراءات المقارنة وفقاً للشروط المتضمنة في الدالة الرياضية (2). وفي حالة كان الصامت الأول والثاني تنطبق عليه أيًا من الشروط المذكورة، فسيكون التوليد خاطئًا؛ أي أن الجذر لآحن في اللسان العربي. وإلا فستكون النتيجة صحيحة؛ أي أن الجذر مستعل أو يمكن استعماله في اللسان ذاته.

على سبيل المثال، إذا تم إدخال الصوامت "ك ت ب" في مكان الإدخال على البرنامج، فستكون القيمة المعادة "صحيحة"، لأن الحرف الصامت الأول "ك" غير مكرر في الموقع الثاني للجذر، ولا يوجد أي عنصر صامتي من عناصر مجموعته (ق، ج، ك) على مستوى الموقع الثاني للجذر، والأمر نفسه بالنسبة للصامت (ت) الذي لم يخرق، بدوره أي قانون من قوانين التأليف بين الجذور الثلاثية في اللسان العربي المعيار.

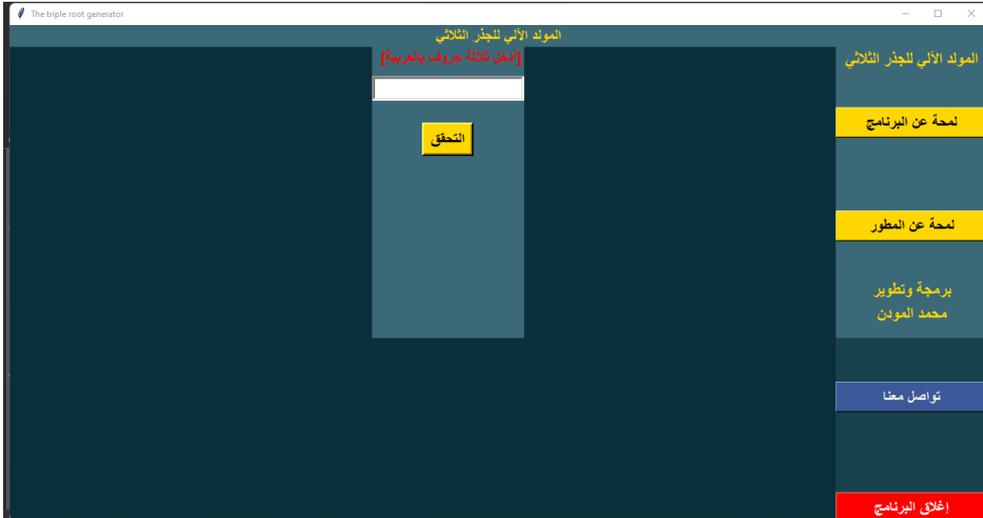
وإن بإمكان هذه الدالة أن تسعفنا في الشروع لبلورة برنامج لتوليد الجذر الثلاثي في اللسان العربي، عن طريق تحويل هذه الدالة إلى أكواد برمجية، وهو ما يؤكد مدى أهمية القوانين الصوتية والصوتية في توليد المفردة العربية.

3. توصيف المولد الآلي للجذور الثلاثية في اللسان العربي

اشتغلنا على توظيف لغة بايثون (Python) باعتبارها لغة برمجية عالية المستوى تتميز باعتمادها على أقل قدر ممكن من الشيفرات، بحث يمكننا اختزال عمليات معقدة في سطر برمجي واحد؛ لذلك نجد لها حضورا وازنا في تطوير الويب، وعلوم البيانات والذكاء الاصطناعي، ومعالجة اللغات الطبيعية.

وسعينا إلى ترجمة الدالة الرياضية الواردة في (2) إلى أكواد برمجية بلغة بايثون، انطلاقا من التصور المعبر عنه في الخطوات الثلاث الواردة في (2.1) من هذا المقال، فخلصنا إلى برنامج يولد آليا الجذور الثلاثية التي يسمح بها نظام اللسان العربي المعياري (العربية الفصحى)، وذلك بفحص الحروف المدخلة من المستعمل ومدى موافقتها للدالة (2)، دون أن يتضمن البرنامج أي لائحة معجمية، أو قاعدة بيانات مخزّنة للجذور الثلاثية التي تتضمنها المعاجم العربية الورقية أو الإلكترونية. وهذا ما مكننا من بلورة البرنامج الآلي:

(3) صورة للبرنامج الآلي المولد للجذر الثلاثي في اللسان العربي المعياري



تُظهر (3) صورة للبرنامج الآلي المولد للجذور الثلاثية في اللسان العربي المعياري، وتعدّ خانة الإدخال التي تظهر في وسط الصورة أهم ما في البرنامج؛ إذ من خلالها، يُدخل مستعمل

البرنامج ثلاثة حروف، كما تنص على ذلك التعليلة باللون الأحمر "أدخل ثلاثة حروف بالعربية". ثم زر التحقق الذي يتكفل بنقل الحروف المدخلة وعرضها على الدالة الرياضية المعبر عنها في (2)، فإذا احترمت كل شروطها تظهر للمستعمل رسالة تقول: "الجذر صحيح، حسب قوانين هندسة السمات في الصوارة اللاخطية".

أما إذا اختل أي شرط من شروط الدالة (2) فإن ذلك يعني أن الحروف المدخلة لا يمكن أن تولد جذر ثلاثيا للمفردة في اللسان العربي المعيار، وبالتالي تظهر للمستعمل رسالة تقول: "هناك قوانين صوتية تمنع من تكوين هذا الجذر في اللسان العربي المعيار". أما باقي الأزرار فهي تعطي معلومات إضافية عن البرنامج، كلمحة عن البرنامج، أو لمحة عن المبرمج والمطور له، ثم زر للتواصل معنا قصد تلقي ملاحظات المستخدمين، أو للإجابة عن تساؤلاتهم....

ويمكن أن نقدم أمثلة لاستخدام البرنامج مع محاولات اختبار حروف معينة، وهو ما سنُظهره في (4)أ، و(4)ب.

(4) أمثلة من فحص مدى سلامة نتائج البرنامج



يستطيع القارئ أن يلاحظ أن البرنامج اشتغل بنجاح مع الإدخالين معا؛ حيث أعطى نتيجة إيجابية مع إدخال (4) أ الذي قدم فيه المستخدم ثلاثة حروف هي: (ع م ل) فكان الخرج هو أن الجذر صحيح حسب قوانين هندسة السمات في الصوتية اللاخطية. أما الإدخال في (4) ب فكانت نتيجته خاطئة لأن الحروف الثلاثة المدخلة (ض ض ب) تخرق شرطا من شروط الدالة (2)؛ وهو تكرار الصامت (ض) على مستوى الموقع الأول والثاني للجذر الثلاثي. وهو ما لا يسمح به مبدأ (ocp) في الصوتية اللاخطية، كما أسلفنا.

4. الخاتمة:

حاولنا في هذا المقال بناء برنامج يحاكي ما يحدث في دماغ المتكلم العربي ليقوم بتوليد الجذور الثلاثية التي يسمح بها نظام اللسان العربي المعياري، بناء على قواعد صوتية، على شكل دوال رياضية، مستنبطة من دراستنا لآليات تشكيل المفردة العربية، ويمكننا إجمالاً تلخيص أهم ما توصلنا إليه في الآتي:

- إن المعالجة الآلية للمفردة في اللسان العربي تقتضي الإلمام بالنظريات الصرف صوتية ما بعد المعياري؛ لكون التوصيف العربي القديم لآليات تشكيل المفردات غير كاف، لأنه لم تكن دواعي له في زمن التعقيد.
- إن من شأن تطبيق مبادئ الصوتية اللاخطية أن يسفر عن استنباط قواعد صوتية تتحكم في تشكيل الجذور الثلاثية في اللسان العربي المعياري، وأنها قواعد قابلة للصورة.
- توجد دوال تتحكم في تشكيل المفردة العربية انطلاقاً من الكشف عن دالة تشكيل الجذور الثلاثية.
- برمجة برامج صرافية تحاكي ما يحدث في دماغ المتكلم العربي، انطلاقاً من مبدأ اللسانيات التوليدية القاضي بكون عقل الإنسان أشبه بالحاسوب، إذ يعتمد على التوليد بدل الاستدعاء من معاجم مخزنة في الدماغ، من شأنه أن ييسر عملية تعليم العربية وراقمتها، بشكل ينسجم وطبيعة التحديات المعاصرة.

- يمكن بتطبيق مبادئ الصوتية اللاخطية مثل مبدأ المحيط الإجاباري (OCP) أن يسعف في الانتقال إلى حوسبة مستويات أعمق؛ مثل التفكير في العلاقات القائمة بين الجذور والصيغ. وبالتالي حل إشكالات المفردات التي صنفت مع السماع في الدرس اللغوي العربي القديم؛ أي معالجة إشكالات القضايا التي لم تخضع لقوانين الصرف العربي القديم.

قائمة المراجع:

- الغامدي، منصور بن محمد، وآخرون، (2017) مدخل إلى اللسانيات الحاسوبية، مركز الملك عبد الله بن عبد العزيز الدولي، (السعودية).
- التاقي، محمد: (1992) التكرار الصامت والتعاقب الصائتي في اللغة العربية، قضايا في اللسانيات العربية، منشورات كلية الآداب والعلوم الإنسانية، ابن مسيك، الدار البيضاء، صص 144-172.
- المودن، محمد: (2022) إشكالات مصادر الأفعال الثلاثية في الصرف العربي، مجلة إشكالات في اللغة والأدب، جامعة تامنغست، مجلد 11، العدد 1، صص 75-88 الجزائر.
- المودن، محمد: (2022) التمثيل الصوري للمصادر الثلاثية المجردة في العربية المعيار، مجلة اللسانيات التطبيقية، جامعة الجزائر 2، مجلد 6، العدد 3، صص 412-424، الجزائر.
- المودن، محمد، (2022) مشاريع الحوسبة في اللسانيات العربية: مشاريع بنية الكلمة نموذجاً، ضمن كتاب: في اللسانيات العربية: ضمن كتاب قضايا في المنهج وتساؤلات حول المنجز، منشورات ألفا للوثائق، (الجزائر)، ص 527-541.
- دالي، محمد: (2019)، مقدمة في علم الحاسب، جامعة حماة، (سوريا)، صص 131-136.
- المهيوبي، عبد العزيز بن عبد الله: (2018) إشكالية تطوير محلل صرف حاسوبي دقيق للغة العربية، مجلة اللغة العربية وتعليمها للناطقين بغيرها، العدد 22.
- نبيل علي، (1987)، العربية والحاسوب، عالم الفكر، المجلد 18، العدد 3.
- هاري فان درهالست، ونورفال سميت، (1992)، الفونولوجيا التوليدية الحديثة، ترجمة مبارك حنون وأحمد العلوي، منشورات دراسات سال، الدار البيضاء.
- McCarthy, J, A prosodic theory of nonconcatenative morphology, (1981), Linguistic Inquiry, Vol. 12, No. 3.
- McCarthy, J, OCP Effects: Gemination and Antigemination, (1986), Linguistic Inquiry, Vol. 17, No. 2.

الهوامش:

- 1 - ينظر: المودن، محمد، (2022) مشاريع الحوسبة في اللسانيات العربية: مشاريع بنية الكلمة نموذجاً، ضمن كتاب: في اللسانيات العربية: ضمن كتاب قضايا في المنهج وتساؤلات حول المنجز، منشورات ألفا للوثائق، (الجزائر)، ص 527-541.
- 2 - انظر المرجع نفسه.
- 3 - انظر المرجع نفسه
- 4 - المهيوبي، عبد العزيز بن عبد الله: (2018)، إشكالية تطوير محلل صرف حاسوبي دقيق للغة العربية، مجلة اللغة العربية وتعليمها للناطقين بغيرها، العدد 22، ص 56.
- 5 - نبيل علي، (1987)، العربية والحاسوب، عالم الفكر، المجلد 18، العدد 3، ص 73.
- 6 - انظر: المودن، محمد: (2022) إشكالات مصادر الأفعال الثلاثية في الصرف العربي، مجلة إشكالات في اللغة والأدب، جامعة تامنغست، مجلد 11، العدد 1، صص 75-88.
- 7 - انظر: المودن، محمد: (2022) التمثيل الصوري للمصادر الثلاثية المجردة في العربية المعيار، مجلة اللسانيات التطبيقية، جامعة الجزائر 2، مجلد 6، العدد 3، صص 424-412.
- 8 - انظر: المودن، محمد، (2022)، مشاريع الحوسبة في اللسانيات العربية: صص 527-541.
- 9 - الغامدي، منصور بن محمد، وآخرون، (2017) مدخل إلى اللسانيات الحاسوبية، مركز الملك عبد الله بن عبد العزيز الدولي، (السعودية). ص 63.
- 10 - see: McCarthy, J, OCP Effects: Gemination and Antigemination, (1986), Linguistic Inquiry, Vol. 17, No. 2, p209.
- 11 - ينظر: التاقي، محمد، (1992)، التكرار الصامت والتعاقب الصائتي في اللغة العربية، قضايا في اللسانيات العربية، منشورات كلية الآداب والعلوم الإنسانية، ابن مسيك، الدار البيضاء، صص 172-144.
- 12 - إلى جانب نظريات صوتية منها:
- الصوتة العروضية (Metrical phonology) مع برينس وليبيان (1977) Liberman & Prince وهايس (1980) Hayes و كاجر (1989) Kager ...
- الصوتة المعجمية (Lexical phonology) مع كيبارسكي (1985) Kiparsky ، هالي وموهنان (1985) Halle & Mohanan
- الصوتة المقطعية (Syllabic phonology) مع كايزر وكليمنتس (1983) Clements & Keyser، كاي ولوفينستام (1984) Kaye & Lovenstamm.
- الصوتة التطريزية (Prosodic phonology) مع سيلكرك (1984-1986) (...)
- هندسة السمات الصوتية (Geometry of Phonological Features) مع كليمنتس (1985) Clements وساجي (1986) Sagey.

- النظرية الأمثلية (Optimality Theory) مع برينس وسمولانسكي (1993) Prince & Smolensky وسمولانسكي وليجيندر (2006) Smolensky & Legendre.
- 13 - هاري فان درهالست، ونورفال سميت، (1992)، الفونولوجيا التوليدية الحديثة، ترجمة مبارك حنون وأحمد العلوي، منشورات دراسات سال، الدار البيضاء، ص 11.
- 14 - لمزيد من التوضيح انظر: المودن، محمد. (2022). التمثيل الصوري للمصادر الثالنية المجردة في العربية المعيار. مجلة اللسانيات التطبيقية (3)، الصفحات 412-424.
- 15 - see: McCarthy, J, A prosodic theory of nonconcatenative morphology, (1981), Linguistic Inquiry, Vol. 12, No. 3, p382.
- 16 - ينظر: دالي، محمد: 2019 مقدمة في علم الحاسب، جامعة حماة، (سوريا)، صص 131-136.