مجلة العمدة في اللسانيات وتحليل الخطاب
Issn:2572-0058/E-issn:1969-2676
https://www.asjp.cerist.dz/en/PresentationRevue/485

المجلد:06 العدد: 02 (2022) ص391/398

حوسبة اللّغة العربيّة في ظلّ الإشكالات التّقنيّة والمُهَيِّئات العلميّة

Automated processing of Arabic under technical and scientific problems

محمّد حرّاث*

جامعة حسيبة بن بوعلي الشّلف – الجزائر-

momo.adab@gmail.com

معلومات المقال

تاريخ الارسال: 2022/01/05

تاريخ القبول: 2021/02/24

الكلمات المفتاحية:
- ✓ اللسانيات الحاسوبية
- ✓ أنظمة الحاسوب
- ✓ اللّغة العربيّة
- ✓ الاشتقاق
- ✓ الخوارزميات

الملخص:

تتحدث المداخلة: عن محورين مهمين في العلاقة بين خصائص اللغة العربية وأنظمة الحاسوب التقنية، فالناظر لواقع العربية حاسوبيا يجد أنها تواجه معوّقات عدّة تقف في طريق حوسبتها؛ لأنّ الحاسوب أول ما أُنشئ إنما أنشئت أنظمته بناء على متطلبات اللغات الغربية وخصائصها، لذا فمن أجل تطويع أنظمة الحاسوب حتى تتواءم مع خصائص العربية يحتاج ذلك اشتغالا دؤوبا من قبل علماء الحاسوب واللسانيين؛ وهو ما تنوء به اللسانيات الحاسوبية في العصر الراهن.

وفي المقابل، وعلى الرغم من هذه الإشكالات والعوائق، فإنّ اللغة العربية تتميز عن غيرها من اللغات بخصائص داخلية ومهيّئاتٍ علمية، تجعل من العربية مناسبة لدخولها نظام الحاسوب، ذلك أنّ بين خصائص العربية وأنظمة الحاسوب كثيرا من نقاط التشابه والتلاؤم؛ فهي لغة رياضية اشتقاقية تخضع لخوارزميات مضبوطة..

Article info

Received *05/10/2021*

Accepted *24/02/2022*

Keywords:
- ✓ *computer linguistics*
- ✓ *computer system*
- ✓ *Arabic*
- ✓ *Derivation*
- ✓ *Algorithms*

Abstract :

He Arabic language suffers from certain problems that hinder its automatic processing; Because computer systems were created on the basis of the requirements and characteristics of foreign languages, in order to adapt computer systems to the characteristics of Arabic, this requires a constant work of computer scientists and linguists; this is what computer linguistics endures in the current era.

On the other hand, despite these problems and obstacles, Arabic is characterized by internal and scientific characteristics, which make Arabic suitable for entry into the computer system. It is a derived mathematical language subject to exact algorithms.

. مقدمة:

ممّا هو معروف أنّ للعربيّة إشكالات حاسوبية، لا تزال للأسف قائمة إلى يومنا هذا، فالحاسوب لا يزال منحازًا إلى اللغة الإنجليزية. حتّى إنّ معظم الأبحاث في حوسبة العربيّة ورقمنتها، قام بها باحثون غير عرب، وهم ينظرون إلى العربية بعين لغتهم. دون أن ننسى كذلك عوائق التّشكيل، وكذا إشكالية الفهم الآلي للغة العربية، وهو حلمنا جميعا، والذي يحتاج إلى مَكْنَزٍ يحوي قاعدة بيانات إحصائيّة، ثمّ إلى برامج سياقيّة دلاليّة تصنيفيّة. وكذا ما يسمى بالذكاء الصناعي.

مع العلم أنّ الحاسوب قد يساعدنا كثيرا في خدمة العربيّة، وإنّ هناك حلولا كثيرة، لا تزال افتراضا مطروحا، ينتظر أن يصبح واقعا معاشا، فالحاسوب يمكنه –مثلا- مساعدتنا في إنشاء: (برامج تعليميّة، وبرامج للصّرف الآلي، وبرامج للنحو والإعراب، وبرامج للنّطق ضمن خصائص الصّوتيات العربيّة، وبرامج للتّرجمة، وبرامج للتّشكيل الآلي). وهذا ما يضطلع به علم اللسانيات الحاسوبي، الذي يتصدّى إلى إشكالات تطبيق اللغة العربية حاسوبيا. وقد تحقق بعض هذا الذي تقدم، ولكن يبقى محتاجا إلى بعض الترشيد والتوجيه.

ومما يقلّل من حدّة هذه الإشكالات والمعوّقات، ومما يشجّع على تجاوزها هو ما تمتاز به اللغة العربية عن غيرها من اللغات بخصائص علمية تجعلها مُهيأة لدخولها أنظمة الحاسوب، من أجل حوسبتها ورقمنتها؛ ومن هذه الخصائص: الجذر، والأوزان الصرفية، والإيقاع، والاشتقاق، وغيرها من المهيئات العلمية.

1- إشكاليات العربية في الحاسوب:

تمتلك اللغة العربيّة من المؤهّلات ما يجعلها تتأقلم مع أنظمة الحاسوب التي تتألف من "منظومةٍ برمجيّة منطقيّة، قوامها الخوارزميات الصّارمة، التي لا تشتغل بالظّنّ أو بالنّسبة، ولذلك فإنّ القواعد التي يجبُ أن تُصاغَ لهذه الغاية، يجب أن تكون صوريّةً وحاسمةً، لا تقبل أكثر من تأويل واحدٍ لكلّ قضيّة"[1]. وهذا من أعقد الأمور وأعسرها بالنّسبة للّغة العربية، ذات الدّلالات الواسعة.

فأهم إشكالية قد تواجه العربية دون غيرها من اللغات، هي الاتساع الدلالي الكبير للمعاني، فللعربية تعامل خاص مع المعاني، حتى إن ظهور علم البلاغة أو علم الإعجاز كان للإحاطة وسبر أغوار هذه المعاني والدلالات، فالعربي قد يعبر عن المعنى بضده، مبالغا في التوكيد، وقد يعبر عن اليقين بالشك، وعن الشيء بضده، فيورد الكلام ولا يريد معناه الأصلي، فظهر التضاد والتورية والمجاز والاستعارات، فهي وإن كانت في بعض اللغات موجودة ، فإنها بالتأكيد ليست بالصورة التي عليه في العربية.

وهذا ما يقف عائقا أمام مسألة الترجمة الآلية للغة العربية، فحين يُدخل الحاسوبيُّ الكلمة كمعطى من معطيات المِنْطاق المترجِم، يضع أمام الكلمة ما تحتمله من معان، لكن عند الاستعمال، يكون السياق وحده هو من يحدد المعنى المقصود لتلك الكلمة، من بين غيرها من المعاني، التي كان يمكن أيضا أن تكون مقصودة. فهذا الإشكال موجود في اللغات، لكنه بحدّة أقل، وهو بخطورة أكبر موجود في العربية؛ هذه اللغة التي كثيرا ما تلجأ إلى المجاز والاستعارة.

وعلى الرّغم من أنّ التّطبيق على اللّغة العربيّة ما زال في بداية الطّريق، فإنّ ما تتميّز به اللّغة العربيّة عن غيرها يجب أن يدفع بها إلى مقدّمة اللّغات العالميّة في التّعامل مع الآلة، فهي تقوم على الجذر والوزن، وهما مكوّنان رياضيان، لا يوجدان في أغلب اللّغات[2]. فالجذر يتولّى وضع البنية الأساسيّة للكلمة، ويتولّى الوزن وضع هيكلها العام، وكلّ هذا يستدعي تضافرا ضروريا بين علماء التّخصّصين: علماء اللّغة وعلماء الحاسوب.

وهناك من ينفي إمكانيّة أن تتطوّر العربيّة حاسوبيّا، مستدلا بثنائية الإنتاج والاستهلاك. فالإشكال يكمن في موقع العربيّة الاستهلاكي، "فالمنتج يستعمل لغته، ويطبّق عليها، ويعالج المشاكل انطلاقا من خصوصيّات لغته؛ أمّا المستهلك فيضطرّ إلى استعمال الألسن الأخرى، ويبقى لاهثا وراء نسق التّقدّم العلميّ الذي يحقّقه المبتكر المصنِّعُ والمنتج، وهنا بالذّات يكمن مشكل المعالجة الآلية العربيّة، التي لا ترقى إلى المستوى العالمي رغم الجهود المبذولة"[3]. ومن هنا، لمّا كان الحاسوبُ قد صُمّمَ طبقا للغة المنتِج (الانجليزية)، فمن الصّعب إعادة ملاءمته مع طبيعة العربية، إلا بمساعدة المنتج على ذلك.

ومن الإشكالات التي واجهت العربية نذكر:

1-1- الوصف والتّوصيف:

من إشكالات العربيّة في الحاسوب إشكاليّة توصيفها، وقد فرّق نهاد الموسى بين هذين المصطلحين، فقد سمّى ما يُعمل للإنسان وصفًا، وما يُعمَلُ للحاسوب توصيفًا. فوصفُ اللّغة عند العرب كان جزءٌ منه مبنيّا على المستَقبِل، فيساهِمُ بحدسِهِ وسليقته وخبرته في فهم الكلام، أمّا توصيف الكلام مجرّدًا عن هذه العناصر فلا يوصِلُ إلى المعنى، فالحاسوبُ يفتقر إلى هذا العنصر البشريّ الخالص، فلو قلنا: (دخل موسى المستشفى) فهمها الإنسان، ولو قلنا: (دخل المستشفى موسى) أدرك الإنسانُ أنّ موسى هو الفاعل وإنْ كان متأخّرًا، وأنّ المستشفى لا يمكن أن يكون فاعلا. أمّا الحاسوب فيحتاجُ إلى قائمةٍ من البيانات المعجميّة حول كلّ كلمةٍ تحدّدُ سلوكها في الجملة، حتى يدرك أنّ (المستشفى) لا يدخل (موسى).

وكذلك إذا سمع الإنسان جملة: (رأى الحصان السّوط) و(سمع الحصان الصّوت) أدرك أنّ المنطوق الأوّل (السّوط) غير المنطوق الثّاني (الصّوت)، وإنْ كانا متماهيين في السّمع، أمّا الحاسوب فإنّه يحتاج دليلا مستقصًى من الدّوال، لإقامة الفرق، وإدراك المعنى المراد من الجملتين[4].

1-2- السّياق/ المقام:

مع كلّ هذا التّطوّر في تجريد اللّغة في العقل، وتوصيفها للحاسوب، فإنّ السّياق يبقى مشكلةً عالقةً، أو مُرجَأَةً إلى حين، وإنْ كان هناكَ محاولاتٌ لتحديده وضبطه، إذْ ليس له ضوابط تضبطه، كالنّحو والصّرف، كما لا يمكننا رصد جميع السّياقات التي يتعرّض إليها الإنسان في عملياته الكلاميّة[5].

1-3- الكفاية:

وتقوم على أربع ركائز:

أوّلا: محاولة إدخال قواعد اللّغة العربيّة، بنظامها الصّوتي، وأنساقها الصّرفيّة، ونَظْمِها الجُمَليّ، وأنحاءِ أعاريبها، ودلالات ألفاظها، ووجوه استعمالاتها، وأساليبها البيانيّة، وأحكام رسمها الإملائيّ؛ إلى نظام الحاسوب.

ثانيا: إنتاج ما لا يتناهى من الأداءات اللّغويّة الصّحيحة، بناءً على معطيات أوّليّة، وهو ما نسمّيه بالقياس والتّمثيل.

ثالثا: مرجِع لتمييز الصواب من الخطأ.

رابعا: استخدام اللّغة وفقا لسياقات ومقتضيات.

وكلّ هذه الرّكائز هي التي يعتمدها الإنسان العربيّ في نطقه وكلامه، وتحقيق الكفاية يكون في إدراجها على الحاسوب، وهنا الإشكال، فهذه الكفاية تتحقق للإنسان بالسّليقة، وتُهَذَّبُ بالمِران، والدُّربة، والتّثقيف. فكيف يكون للحاسوب، وهو جهازٌ أصمُّ، مثلما كان للإنسان[6]؟ إنْ لم يكن هذا مستحيلا، فصعب جدا على الأقل.

1-4- الخط:

واجه الخطّ العربيّ حاسوبيا إشكالات عدة، وصعوبات كثيرة؛ منها:

-لوحة المفاتيح التي كانت مهيّأة أصلا لنظام اللّغة الإنجليزية.

-اتجاه الخطّ العربي من اليمين إلى اليسار، عكس الإنجليزية.

-تعدّد أشكال الحرف العربي بحسب موقعه في كلّ كلمة، كالهمزة.

-اتّصال الحروف وتلاصقها ببعضها، عكس حروف الإنجليزية المتفرقة المتقطّعة.

-إشكالية غياب الشّكل وعلامات الضّبط.

لكن مع تحدّي هذه الصّعوبات، تَمّ التّغلّب عليها، وتحقيق حلم إدخال الحرف العربي إلى الحاسوب، حتى وصل الأمر إلى تعرّف الحاسوب على الخطوط اليدويّة والإمضاءات[7]. وأفاد محمد مراياتي أن ثَمَّ إشكالية تستدعي مضاعفة الجهود؛ لوضع تعامل الحرف العربي مع الأجهزة والمعدات؛ مثل:

-ترميز الحرف المكتوب وتقييسه.

-ترميز الحرف المنطوق.

-توزع الحروف العربية على لوحة الملامس.

-تقييس الأقلام العربية وإظهارها على الشاشات والطابعات.

-تحرير النصوص وتنضيدها.

-معاملة الحرف العربي على شبكات الاتصال من حيث نقل المعلومات أو أمنها.

-ضغط النصوص العربية؛ بغية تخزينها في ذاكرة الحاسوب اقتصادياً.

-تحاور المعوقين مع الآلة باللغة العربية[8].

1-5- الدّلالة اللّغويّة آليا:

لا يمكن معالجة النّصوص دون أن تتوفّر معلومات كافيةً عن دلالات الألفاظ المكوّنة لتلك النّصوص، ومن هنا اكتست الدّلالة أهمّيّة عند علماء اللّسانيّات الحاسوبية.

وتحتاج المعالجة الدلالية الآلية للغة العربية مقدارًا كبيرًا من المعلومات عن مختلف جوانب اللغة. وهذه المعلومات يجب أن تكون مرتبة ومبوبة بنسق معين. وليس هناك أفضل من ترتيب هذه المعلومات من قواعد البيانات. وهذه المعلومات تشمل المعلومات عن الألفاظ وما تحتويه من دلالات مختلفة، وصرف وما يجوز أن تستعمل معه الكلمة وما لا يجوز، والكلمات المقاربة والمضادة، وكل ما يوصل إلى دقة دلالة الكلمة من وسائل[9].

2- المهيّئات العلميّة لتقننة اللّغة العربيّة حاسوبيا:

2-1- الثنائيات:

إنّ الناظر الممعن إلى مخصصات هذه اللغة، يستشف أنّها تملك من المبررات العلمية ما يجعلها تتبوء أعلى مقام في فلك الرقمنة والتقانة. كيف لا وهي لغة طبيعية؛ أي تتماشى وقانون الطبيعة. فإذْ علمنا أنّ الله تعالى خلق من كل شيء زوجين اثنين،

وأنّ الكون قائم على هذه الثنائية الزوجية، فنعلم أنّ لغة الحاسوب حين تأسس قائمة على اللغة الثنائية (01). وفي المقابل نجد العربية قائمة على هذا النظام الثنائي المتمثل: في الحركة والسكون أو في الصوامت والصوائت. وهذا ما يجعل العربية مُهيأة وجديرة بِتَقْنَنَتِهَا ورَقْمَنَتِها.

فلا يخلو الحرف في العربية من أن يكون متحركا: (مرفوعا أو منصوبا أو مجرورا). أو يكون في المقابل ساكنا. وقد اصطلح العلماء قديما على المتحرك والساكن بالرمزين: (/0) فالخط يقابل الحرف المتحرك، والدائرة تقابل الحرف الساكن. وهذه الثنائية الزوجية تقابلها لغة الحاسوب الثنائية أيضا: (1-0).

ومن الثنائيات التي تتميز بها العربية أيضا، ثنائية الصوامت الصوائت. فالحروف العربية هي في الأصل مقاطع صوتية غير قابلة للتقطيع؛ أي أنّ الحرف هو أصغر وحدة صوتية. فالصوامت **(Consonants)** هي الحروف الأقل وضوحا في السمع، أو الحروف المنخفضة في الصوت؛ نحو: (ب، ت، ث ...). وأمّا الصوائت **(Vowels)** فهي تلك الحركات: من ضمة وفتحة وكسرة. وكذلك: الواو والألف والياء. فالأولى: صوائت قصيرة، والأخرى صوائت طويلة. وهي أوضح في السمع، وأعلى في الصوت.

2-2- الجذر:

أضف إليها من المُهيّئات أنها لغة قائمة على الجذر، من الثنائي إلى السداسي. فالجذر صيغة أو قالب شكلي تصاغ منه كل الكلمات وتصريفاتها، وما يطرأ عليها من زيادات وإعلالات وإبدالات. فمن خلال الجذر تستطيع المناطيق الرقمية للحاسوب أن تسيطر على الصيغة الأصلية للكلمة قبل تفرّعاتها، فيسهل التحكم في الكلمة وتصاريفها. وهذا مهيّئٌ ثان من مهيئات العربية واستعدادها الطبيعي لدخول عالم الرقمنة من بابه الواسع.

فالجذر (ك ت ب) له صرفيا ولغويا احتمالات محدودة، من حيث الترتيب؛ أي ترتيب الحروف، ومن حيث تشكيل الحروف. لذا يعطى للمنطاق هذا الجذر، فيعطي المنطاق عبر عملية رياضية معقدة، جميع الاحتمالات الممكنة: المستعملة والمهملة؛ مثل نظام التقليبات الذي استعمله الخليل بن أحمد الفراهيدي في معجمه العين. وما على اللساني الحاسوبي إلا تصنيف هذه الاحتمالات، وهذا ما يسمى التوصيف؛ أي توصيف اللغة حاسوبيا.

فبناء على خِصّيصة الجذر، نستطيع السيطرة على جميع كلمات العربية بطريقة رياضية حاسوبية سريعة وسهلة.

2-3- الميزان الصرفي: كذلك من المهيئات أنّ العربية لغة قائمة على الوزن، أو الميزان الصرفي، وهو قالب معنوي إيقاعي، يتحكم في التدفق الصوتي للكلمات، وهي أبنية صرفية تصاغ على منوالها كل كلمة عربية، فلا تكاد تخرج كلمة عربية فصيحة

عن ميزان صرفي معين، ويسهل ضبط هذه الموازين على مناطيق الحواسيب وبرامجه، القياسي منها، وحتى السماعي، الذي على الرغم من حرية حركيّته إلى أنه مضبوط أيضا. فهذا الميزان الصرفي يهيئ العربية لرقمنتها في الحاسوب.

ويقابل الميزان الصرفي موازين أخرى: كالميزان التصغيري، والميزان العروضي، والميزان المقطعي. ويبقى الميزان الصرفي (فعل) في جذره الثلاثي هو الضابط لكل الموازين الصرفية الأخرى، لأنّ أكثر كلمات العربية ميزانها الصرفي ثلاثي. لذا بُنيت على الميزان الثلاثي بقية الموازين غير الثلاثية الأخرى.

2-4- مهيّئات أخرى:

بل إن من مخصصات العربية أنّ لتراتب الحروف ومواقعها داخل الكلمة ضوابط، فبعض الحروف لا تلتقي بحروف أخرى في كلمة واحدة، وإلا كانت الكلمة أعجمية أجنبية، وبعض الحروف يجب حضور أحدها في كل كلمة، وما غاب منها هذه الحروف فهي كلمة أعجمية دخيلة غير فصيحة. فوضع هذه الاحتمالات الرياضية يساعد في وضع منطاق يحلل الكلمات ويدرجها ضمن الفصيح الأصيل، أو غير الفصيح الدخيل.

فمن ذلك، تمثيلا، أنّ كل لفظة تبدأ بالنون، ثمّ تليها الراء ليست عربية فصيحة؛ أو أن يكون آخر اللفظة زايا مسبوقا بدال، أو أن تجتمع في لفظة واحدة الصاد مع الجيم، أو الجيم مع القاف، أو الياء والتاء والسين، أو الجيم والطاء؛ نحو: الصولجان، والمنجنيق، والبستان. فيمكن إعطاء هذه الأوامر للحاسوب، ليقوم بتحليل ومعالجة كل لفظة مُدخلة ليصنفها في العربي الفصيح، أو في المعرّب والدّخيل والمولّد.

دون أن ننسى كذلك كثيرا من المخصِّصات والمهيئات كمخارج الحروف وصفاتها، التي تم ضبطها بحكمة. كل هذه المهيئات التي سبق ذكرها تبين لنا بما لا يدع للشك مجالا أن اللغة العربية لغة علمية طبيعية رياضية رقمية تقنية.

خاتمة:

يبقى في الأخير أن نفتح بابا للأمل، بعد أن تكللت الكثير من الجهود الحاسوبية بالنجاح، فنجد على الشابكة الكثير من البرامج والمناطيق العربية، التي تخدم لغتنا، ولم يبق من هذه الإشكالات التي تقدم ذكرها إلا القليل، فإذا نظرنا أولا إلى الجهود التي بذلت وتكلّلت بالنجاح، ثم نظرنا ثانيا إلى المهيئات الطبيعية العلمية التي تمتاز بها اللغة العربية عن غيرها من اللغات، فإن ذلك يجب أن يكون حافزا وداعما لمواصلة السير بالعربية نحو الحوسبة التامة الشاملة لجميع نظمها، وأبوابها، ومستوياتها.

الهوامش:

[1] - الحنّاش، محمد، (2002م)، اللّغة العربية والحاسوب (قراءة سريعة في الهندسة اللّسانيّة العربيّة، أو مقاربة في محاكاة الدّماغ العربيّ لغويا) محاضرة بقسم اللّغة العربيّة وآدابها، جامعة الإمارات العربية المتحدة، ص2.

[2] - المرجع نفسه، ص3.

[3] - حافظ، حافظ إسماعيلي، و العناتي، وليد أحمد، أسئلة اللّغة أسئلة اللّسانيات، ص79.

[4] - العناتي، وليد أحمد، اللسانيات الحاسوبية العربية رؤية ثقافية، مقال متاح على الأنترنت.

[5] - المرجع نفسه.

[6] - المرجع نفسه.

[7] - المرجع نفسه.

[8] - أبو هيف، عبد الله، (2004)، مستقبل اللغة العربية: حوسبة المعجم العربي ومشكلاته اللغوية والتقنية أنموذجا، سورية، مجلّة التراث العربي، العدد93-94، ص97.

[9] - خضر، محمد زكي، (2004)، نحو معالجة الدّلالة في اللغة العربية عبر قواعد البيانات: دراسة أولية لنص القرآن الكريم، المكلة العربية السعودية، جامعة الملك عبد العزيز.