

توظيف الذكاء المحوسب لإنشاء المدونة المحوسبة لقاموس لسان العرب وسبر مادته اللغوية واللسانية والمعرفية

حسن مظفر الرزو

رئيس باحثين أقدم

المعهد العالمي لحوسبة القرآن الكريم والعلوم الإسلامية

العراق

halrizzo@gmail.com

المحور: محاور اللغة العربية في المعرفة والتقنية والتطبيقات والصناعات الحديثة والذكاء الاصطناعي

المستخلص

أسهم الحضور اللافت لآليات المعالجات الذكية التي تستثمر القدرات الهائلة للحواسيب، في لفت انتباه العاملين في مجال اللسانيات المحوسبة الى توظيف هذه القدرات في تثوير محتوى المعاجم والقواميس، وإعادة تشكيلها وفق أنماط مستحدثة، لتوليد بيانات تفاعلية تتسم بانفتاحها على علوم اللغة العربية بجميع مجالاتها، نحواً، وصرفاً، ودلالة، وسياًقاً من خلال استثمار المعالجات الذكية التي باتت قادرة على ممارسة جميع أدبيات علوم اللغة العربية دفعة واحدة على النص، لتعود الى ممارسة سلسلة من عمليات إعادة تشكيل مفردات هذه العلوم، وقواعدها، وشواهداها، لتودعها في بيئة تفاعلية ذكية يمكن استثمارها في توليد نصوص جديدة، أو الإجابة عن التساؤلات، أو تثوير المحتوى المعرفي للمادة العلمية الخصبة التي تزخر بها معاجم لغتنا العربية وقواميسها.

استهدف البحث معالجة محتوى معجم لسان العرب لابن منظور بواسطة خوارزميات الذكاء المحوسب لإعادة هيكلتها، وفق منظور معالجات اللغة الطبيعية باستخدام حزمة من الخوارزميات الذكية التي ستمارس عمليات التحاليل النحوي، والصرفي والدلالي المحوسب لمحتوى المعجم، وتحويلها الى فضاء معرفي. محوسب مفتوح يتيح للمستخدم تثوير المادة المعرفية المودعة في هذا العمل الجليل.

وحاول البحث الإجابة عن طبيعة الفرص المتاحة أمام معالجات الذكاء التوليدي في إعادة هيكلة محتوى المعاجم العربية لإنتاج فضاء دلالي متعدد البعاد يمكن استثمار مادته لممارسة سلسلة متنوعة من عمليات التثوير المعرفي التي ستمنح المستخدم العربي فضاءً معرفياً من نمط جديد.

وقد استخدمت خوارزميات الذكاء المحوسب في تفكيك مادة نص المعجم الى وحدات أولية، ثم مورست عمليات التحليل الصرفي، والنحوي، والدلالي المحوسب على مادة النص مع وسم الكيانات اللغوية التي استوطنت النص (أعلاماً، وأماكن وبلدان) مع إجراء تحليل دلالي محوسب لمسألة حضور الألوان في نص المعجم، والفضاءات الدلالية التي حضرت من خلالها في مادته اللغوية لبيان الفرص المتاحة أمام إنشاء مستويات متعددة من الفضاءات الدلالية التي يمكن أن ينهل من لبابها المعرفي المستخدمون بمختلف توجهاتهم ومستوياتهم المعرفية.

وأظهرت المعالجات الذكية وجود أكثر من فرصة لإعادة تشكيل مادة المحتوى اللغوي واللساني الذي أودع في المعاجم وقواميس اللغة العربية لإنتاج فضاء معرفي محوسب، متعدد المستويات يمكن أن يستثمر كمورد خصب على صعيد الدراسات النحوية، والصرفية، والدلالية. مع إمكانية استثمار مادتها في إنشاء نماذج لغوية ضخمة LLM لدعم نماذج الذكاء المحوسب التوليدي التي ستتصاع لثوابت وقواعد لغتنا العربية.

1. مقدمة:

أوجب التحالف الجديد بين اللسانيات اللغوية والحوسبة الذكية مسألة إعادة تشكيل النص التقليدي بأدوات معالجات اللغة الطبيعية لتجاوز عقبة تمثيل المعرفة البشرية المودعة في النصوص وتضمينها ضمن القوالب التي باتت تفرضها مختلف أنماط المعالجات المحوسبة الذكية. لقد حصلت طفرة مفاهيمية بالتعامل مع مفردات النص نتيجة للخطاظة المعرفية الجديدة New Paradigm التي تشكلت من تكامل المعالجات الذكية، مع أدوات اللغة الطبيعية، والمبادئ التي أرستها اللسانيات المحوسبة خلال السنوات الخيرة بعد ولادة الذكاء التوليدي Generative A.I. فبرزت من رحم هذه الخطاظة المعرفية مفاهيم جديدة، سلسلة من معالجات تفكيك مفردات النص التقليدي الى وحدات أولية أطلق عليها اصطلاح Token، والتي أضحت مادة خصبة تتعامل معها الخوارزميات الذكية، بالسبر والتحليل، للكشف عن الأنماط النحوية، والصرفية، والدلالية التي أودعت بمفردات النص، والعلاقات التي تربط فيما بينها ضمن السياقات التي وردت فيها هذه المفردات. وشهدنا ولادة المدونات المحوسبة التي استمدت ذخيرتها اللغوية من نواتج معالجات اللغة الطبيعية للنص، التي باتت توفر لنا كماً هائلاً غير مسبوق من مادة التحليل المحوسب التي باتت تعالج كل مفردة من مفردات النص بأنماط لسانية متعددة، تبالغ في حراثة المعاني، والدلالات، وتشكيل شبكات المفاهيم، على التوازي مع التحليل النحوي، والصرفي، والدلالي لكل مفردة من مفردات النص وبزمن قياسي غير مسبوق.

ولقد سعينا في هذه الدراسة الى توظيف حزمة من المعالجات اللسانية الذكية لتحليل محتوى احد اهم معاجم اللغة العربية ومدوناتها اللغوية، ألا وهو معجم لسان العرب لابن منظور، فجعلنا مادته مورداً لسلسلة من المعالجات المحوسبة الذكية التي قامت بتكديك النص الى وحدات أولية، تم تشكيلها في مدونة محوسبة لمادة المعجم، ثم باشرنا سلسلة من التحليلات اللسانية الذكية على المحتوى لكي تتحول الى نسق معرفي جديد، حاولنا أن نعرض بعض الفرص التي يمكن أن يوفرها للعاملين في مجال العلوم اللسانية للغتنا العربية، والتتقير في أعماق أكبر داخل الخطاب اللغوي المودع في هذا المعجم الثري، واستنطاق مفرداته في مجال حضور الألوان ودلالاتها اللغوية في هذا المعجم بطريق محوسبة نأمل أن تسهم في فتح آفاق جديدة لكي ننهل من المعاني والدلالات المودعة في مفردات لغتنا العربية الخصبة.

2. معجم لسان العرب: إعادة تشكيل المحتوى:

لسان العرب معجم لا يشق له غبار في دائرة معاجم اللغة العربية، صنّفه اللغوي الكبير محمد بن مكرم الأنصاري الذي عرف بابن منظور الأفريقي¹، جامعاً ذخيرته اللغوية من معاجم كثيرة منها: التهذيب، والمحكم، والصاح، والجمهرة، والنهاية، وحاشية الصاح. وقد انكب على تجويد مادته ورتبه ترتيب الصحاح².

ولقد استثمرنا محتوى النسخة الرقمية لمعجم لسان العرب التي توفرت أكثر من نسخة منها على مواقع متعددة من فضاء الإنترنت لكي نباشر على محتواها معالجات اللغة الطبيعية المحوسبة³. ابتكرت معالجات اللغة الطبيعية المحوسبة (Natural Language Processing (NLP) لتحقيق آمال العاملين في مجال اللسانيات المحوسبة والذكاء الاصطناعي لتمنح الحواسيب فرصة سبر الخطاب البشري، وتفسير معانيه، وإنتاج سياق الخطاب بطريقة تقارب محتوى الخطاب التواصلية الذي يستخدمه الإنسان المعاصر، وذلك عن طريق تجسير الفجوة القائمة بين اللغة والخوارزميات المحوسبة التي تستخدم في ميادين الذكاء الاصطناعي Artificial Intelligence والذكاء التوليدي Generative A.I.

بصورة عامة، تتضمن عملية استثمار معالجات اللغة الطبيعية المحوسبة في إعادة تشكيل محتوى النص التقليدي وهيكله مفرداته بحيث تتوافق مع البيئة الذكية التي تعمل بدائرتها الخوارزميات المحوسبة Computational Algorithms من خلال سلسلة من المعالجات التي تتألف من مجموعة من المراحل.

وتتضمن المرحلة الأولى عملية تفكيك محتوى النص إلى وحدات أولية Tokenization يتم تحديد هويتها من خلال المسافة البادئة التي تفصل بين الكلمة والكلمة التي تليها. ويطلق على كل منها اسم وحدة أولية Token.

ووفق الخطأ المعرفية لمعالجات اللغة الطبيعية فإن الخوارزميات المحوسبة في تعاملها مع هذه الوحدات الولية لا تأخذ بعين الاعتبار البعد المعرفي لحضورها على المستوى النحوي، أو الصرفي، أو الدلالي، وتعتبرها كتلة لسانية صماء ينصب اهتمام الخوارزميات على إحصاء مستويات كراتها بالنص، والموقع الذي تستوطنه في نسيج مادته.

ولتجاوز عقبة كثرة المفردات اللغوية المودعة بالنصوص، والتي تورث الخوارزميات البرمجية، والحواسيب، المزيد من الجهد الرياضي والمنطقي المحوسب يصار إلى استبعاد الوحدات الأولية التي لا تمتلك معنى قائماً بذاتها (كالحروف، وحروف الجر، والضمائر، وغيرها) والتي يطلق عليها اصطلاح كلمات التوقف Stop-Words من قائمة تعدّ سلفاً لاستبقاء المفردات التي تحمل معانٍ، وتؤثر على التحليل المحوسب الدلالي.

تأتي بعد ذلك مرحلة المعالجات اللغوية المحوسبة، والتي تتضمن التحليل النحوي المحوسب للنص، ثم التحليل الصرفي المحوسب، وأخيراً التحليل الدلالي المحوسب. فتمارس الخوارزميات المحوسبة في الشق الأول تحليل وتحديد أقسام الكلام لمحتوى النص، أما في الشق الثاني فتمارس الخوارزميات مهمة تحليل المفردات عن طريق إزالة اللواصق (السوابق واللواحق) لبلوغ جذر الكلمة وجذعها، وإحصاء جميع تفاصيل البيانات الصرفية في النص. أما الشق الدلالي فتمارس الخوارزميات المحوسبة مهمة كشف اللبس الدلالي الذي يصاحب حضور الكلمة الواحدة في سياقات متعددة.

وتعالج مخرجات التحليلات المحوسبة الثلاث بسلسلة من المعالجات الرياضية للكشف عن مستويات جذور الكلمات وجذوعها، وما يرتبط بها من أنماط التي تحصل على الجذر باستخدام الحروف والأصوات لتغيير المعنى وتحديدته والتي تمارس دوراً مهماً في إنتاج المعاني في مختلف مستويات السياقات التي استخدمها المؤلف في إنتاج مادة النص. بعدها بوشر بعملية وسم هوية الكيانات Named Entity Recognition والتي تمارس بواسطة نظم تعلم بواسطة الشبكات العصبونية Neural Networks لوسم أسماء الأعلام، والبلدان، والأماكن، وأية كيانات أخرى لغرض إلحاقها بالمدونة الرقمية واستثمارها في معالجات الذكاء المحوسب.

ثم تأتي بعد ذلك معالجات التمهيد لإنتاج بيانات الذكاء التوليدي حيث تبرز معالجات تضمين الكلمات Word Embeddings لتوليد فضاء دلالي متعدد الأبعاد يحتوي كل بعد من أبعاده الدلالة التي حملتها المفردة في السياقات المتكاثرة التي وردت فيها في الخطاب اللغوي المودع في لسان العرب، والتي تعد من المؤشرات المهمة على الذخيرة اللغوية التي امتلكها ابن منظور من

جهة، وفرص استثمارها في توليد كلمات واصطلاحات مرادفة، ومتلازمات لفظية، مع توفر فرصة سانحة للكشف المعاني من خلال الارتباط الدلالي بين المفردة وبقية المفردات الموجودة في نص لسان العرب، واختصار مقاطع محددة من النص، أو سبر الفضاء الدلالي للمفردات والاصطلاحات المقيمة في نص مدونة لسان العرب.

3. المستخرجات اللسانية من النص المحوسب:

استخدمت البيئة البرمجة للغة بايثون كحاضنة برمجية لمجموع الخوارزميات المحوسبة التي أعدت لممارسة حزمة المعالجات اللسانية تحت مظلة معالجات اللغة الطبيعية بمختلف مراحلها وبدعم من المكتبات البرمجية المتوفرة في هذه البيئة البرمجية الذكية مثل مكتبات: NLTK, STANZA, CAMEL, SpaCy, Gensim & Pytorch، وغيرها من المكتبات التي تحفل بها تطبيقات هذه اللغة الذكية⁴.

واعتمد مبدأ إنشاء مدونة محوسبة⁵ Corpus من مادة لسان العرب لكي تكون مورداً للعاملين في مجال الذكاء المحوسب، بعد أن تعاد هيكلتها بواسطة مجموعة متنوعة من الخوارزميات الذكية التي ستمنحها عمقاً معرفياً من نمط جديد سيتيح لنا فرصة التوغل في أعماق البنيات النحوية، والصرفية، والدلالية لمادة لسان العرب مع وسم هوية الكيانات المقيمة في فضاءه اللغوي، وسبر الفضاء الدلالي لكل كلمة وردت فيه، لكي تكون مورداً معرفياً خصباً، لإنتاج المعاني، واستخلاص المفاهيم⁶.

3.1. مستخرجات تفكيك محتوى لسان العرب:

اعتمدت خوارزمية ذكية لتفكيك محتوى لسان العرب وتحويله الى وحدات لغوية أولية Tokens، وذلك باعتماد وجود الفراغ بين كلمة وأخرى دليلاً للتفرقة بين كل وحدة لغوية وبين الوحدة التي تليها. ثم بوشرت عملية استبعاد كلمات التوقف من الوحدات الأولية نظراً لعدم امتلاكها معنى قائماً بذاتها، كما أنها تشكل عبئاً على موارد الخوارزميات المحوسبة. وقد احتسب عدد تكرارات كل فئة من فئات الوحدات الأولية، والكلمات المستبعدة، وكذلك الوحدات الأولية المتبقية بعد عملية إزالة كلمات التوقف. أنظر الشكل (1).

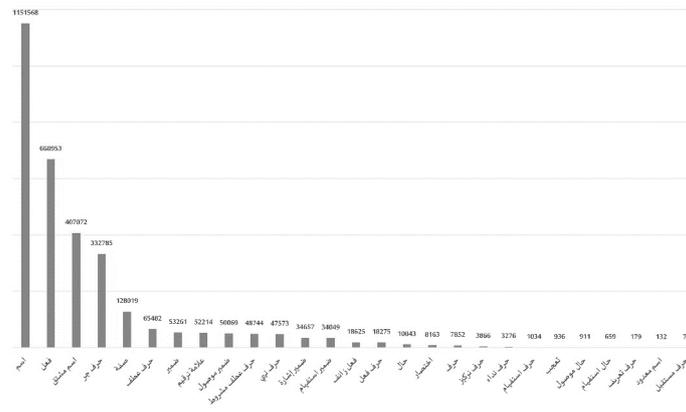


الشكل (1). مادة المدونة المحوسبة لمعجم لسان العرب.

كما يبدو من الشكل أن عدد الوحدات الأولية قد بلغ أكثر من ثلاث ملايين تألف قوامها من 310 آلاف فئة، أما عدد كلمات التوقف فقد بلغ عددها حوالي مليون وحدة، بينما بلغ عدد فئاتها أكثر من ثلاثة آلاف وحدة بقليل.

3. 2. مستخرجات التحليل النحوي والصرفي والدلالي لمحتوى لسان العرب:

بوشرت عملية التحليل النحوي، والصرفي، والدلالي المحوسب لمحتوى معجم لسان العرب بواسطة مجموعة من الخوارزميات البرمجية الذكية. فمورست عملية التحليل النحوي لتحديد أقسام الكلام في الخطاب اللغوي لابن منظور، فتوزعت هذه المفردات على سبعة وعشرين فئة نحوية، تباينت أعداد حضورها بالنص بحسب السياقات التي ناقش فيها ابن منظور المفردات اللغوية في معجمه. أنظر الشكل (2).



الشكل (2). أقسام الكلام التي تم الحصول عليها من التحليل النحوي المحوسب لنص لسان العرب.

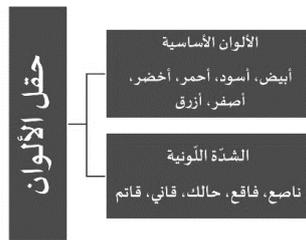
ويبدو واضحاً أن فئة الأسماء، هي الأكثر حضوراً، بينما تأتي بعدها الأسماء المشتقة بالأفعال. أما التحليل الصرفي المحوسب فقد بوشر بواسطة مجموعة أخرى من الخوارزميات البرمجية، التي حرصت على إزالة اللواصق (السوابق واللواحق) من المفردات للظفر بالجذور التي اشتقت منها

والكشف عن المعاني والدلالات المودعة في مادة خطابه، وكشف عن شبكة المفاهيم التي أودعها المصنّف في معجمه.

وتعد مسألة حضور ألفاظ حقل الألوان من المسائل المهمة التي تستأثر باهتمام الباحثين. لأن مفردات الألوان واصطلاحاتها، تتكاثر استخداماتها في وصف كل ما يحيط بنا، حيث نعبر بواسطتها عن معانٍ ودلالاتٍ مختلفة. من أجل هذا فقد اهتم بها أصحاب المعاجم والقواميس فخصّصوا لها حضوراً في معاجمهم للكشف عن الكثير من خصائص ألفاظها، ومشتقاتها، مع بيان دلالاتها الحقيقية والرمزية⁷.

بالمقابل فإن مسألة تعامل المعاجم مع ألفاظ الألوان، ودراسة الصيغ التي صبّت فيها مشتقاتها، ومستويات حضور معانيها، والتباين في دلالاتها، تعدّ من المسائل التي شغلت اهتمام العاملين في مجال اللسانيات والمعاجم⁸. ونظراً للذخيرة اللغوية، والصرفية، والدلالية التي توفرت لنا من خلال المعالجة المحتوى المحوسب لمعجم لسان العرب، فقد وجهنا أنظارنا الى دراسة هذه المسألة مستثمرين الخوارزميات الذكية، وقدرتها على حراثة نص هذا المعجم، للتعقير عن حضور المفردات اللونية فيه، وتحديد مواطنها، ومستويات حضورها، وارتباطاتها مع ألفاظ الألوان المختلفة، أو الألفاظ الأخرى التي تلازم حضورها مع مفردات الألوان.

ولبيان ما يمكن أن توفره لنا المدونة المحوسبة لمعجم لسان العرب، والخوارزميات البرمجية من مجال لساني لدراسة هذه المسألة، والتي ستشدد اهتمام الكثير من الباحثين في حقول معرفية متعددة، للاعتراف من الذخيرة اللغوية والدلالية التي استوطنت في هذا المورد، حاولنا مراجعة أهم المحاور التي وردت فيها ألفاظ الألوان الأساسية ومشتقاتها، مع توجيه الاهتمام نحو الألفاظ التي اهتمت ببيان درجات الألوان وشدتها، وفق معالجة تبنت مفردات النهج المحوسب في عرض نتائجها، تاركين للمتخصصين في مجال حقل الألوان، ومورفولوجيتها، ودلالات ألفاظها فرصة تفسير ما أنتجته المعالجات الذكية من بيانات.



الشكل (5). مجال المعالجات المحوسبة في حقل الألوان بمعجم لسان العرب.

ولاستكمال الشوط في توجيه دفة الخوارزميات المحوسبة في التقدير عن مسائل أخرى تكشف النقاب عن مسألة الحقل اللوني وألفاظه في معجم لسان العرب، سعينا الى الكشف عن ظاهرة المتلازمات والمتصاحبات اللفظية لحقل الألوان في المعجم، وكانت نتيجة المعالجات اللغوية التي مارستها الخوارزميات الذكية، البيانات التي أوردناها في الجدول (3).

الجدول (3) . المتلازمات اللفظية للألوان في معجم لسان العرب.

متلازمات منتخبة تفصح عن تعابير اصطلاحية	التلازم مع لون آخر	عدد المتلازمات	ألفاظ الحقل اللوني
أبيض رطب/ أبيض رأس/ أبيض النعجة/ أبيض الأشغار/ أبيض نابي/ أبيض لهق/ أبيض لمع/ أبيض شيب/ أبيض شعر/ أبيض سادة/ أبيض زرقاء/ أبيض رخماء/ أبيض أصبغ/ أبيض الفرس/ أبيض الطير/ أبيض الضان/ أبيض الشيب/ أبيض الشيء/ أبيض الشاة/ أبيض الروث/ أبيض الزرق/ أبيض الخاصرتين/ أبيض يد/ أبيض يبس/ أبيض ناصية/ أبيض مخالف/ أبيض محور/ أبيض محجل/ أبيض لون/ أبيض قول/ أبيض قاعها/ أبيض غليظ/ أبيض غرة/ أبيض غرباء/ أبيض عيناه/ أبيض عرق/ أبيض عرض/ أبيض ظهر/ أبيض طيب/ أبيض طرف/ أبيض ضخم/ أبيض صقعاء/ أبيض صفة/ أبيض صغيرة/ أبيض شرح/ أبيض شراب/ أبيض شجر/ أبيض سواد/ أبيض ذنب/ أبيض ثرق/ أبيض ذرة/ أبيض داء/ أبيض خصيفة/ أبيض حرف/ أبيض حائط/ أبيض ثمرة/ أبيض بيقة/ أبيض بياض/ أبيض برق/ أبيض أخط/ أبيض أجلح/ أبيض اليد/ أبيض النعمة/ أبيض الموضوع/ أبيض المكان/ أبيض المثابرة/ أبيض اللؤلؤ/ أبيض القياس/ أبيض القطن/ أبيض القاع/ أبيض العيون/ أبيض العصيدة/ أبيض العرب/ أبيض الطريقة/ أبيض الضحك/ أبيض الصحاح/ أبيض الصبغة/ أبيض الشعر/ أبيض الشباب/ أبيض الشاكلة/ أبيض السراويل/ أبيض السبيل/ أبيض الرمة/ أبيض الحلم/ أبيض الحديث/ أبيض الحذقة/ أبيض الجبين/ أبيض الجذوبة/ أبيض الأيام/ أبيض الأرض.	أسود / أصفر / أحمر	580	أبيض .
أسود السواد/ أسود شديد/ أسود الرأس/ أسود العين/ أسود حالك/ أسود الناس/ أسود ببيضاء/ أسود سواد/ أسود الماء/ أسود اللون/ أسود الشعر/ أسود الأرض/ أسود طائر/ أسود شجر/ أسود سواد/ أسود سائح/ أسود جبل/ أسود غير/ أسود المعز/ أسود القوائم/ أسود مشرب/ أسود كفه/ أسود كالليل/ أسود فاجم/ أسود غليظ/ أسود سواد/ أسود بياض/ أسود بطن/ أسود بري/ أسود أروق/ أسود أحوى/ أسود الموت/ أسود المرأة/ أسود الغراب/ أسود الطفتين/ أسود السماء/ أسود الرجلين/ أسود الرجل/ أسود الحياة/ أسود الحلاوة/ أسود الجناح/ أسود الجمع/ أسود الجلدة/ أسود الجبل/ أسود البلوط/ أسود الإبل/ أسود الأذنين/ أسود وجهه/ أسود مطر/ أسود مخضرم/ أسود مخالط/ أسود ماء/ أسود كالسرحان/ أسود قميص/ أسود عود/ أسود عظيم/ أسود ظهره/ أسود سويد/ أسود نداء/ أسود دب/ أسود داء/ أسود خيطم/ أسود خير/ أسود حلية/ أسود حلو/ أسود حرا/ أسود حار/ أسود جلد/ أسود جعد/ أسود ثمر/ أسود برنيق.	أبيض/ أحمر / أخضر/ أصفر	1084	أسود .
أحمر بري/ شجر أحمر/ أحمر شديد/ شعر أحمر/ أحمر الزبح/ أحمر الحديث/ أحمر مصبوع/ رجل أحمر/ أحمر ثمره/ أحمر العين/ أحمر الشيء/ أحمر الدم/ أحمر البيت/ أحمر قاني/ أحمر غليظ/ أحمر الليل/ أحمر اللون/ أحمر القطة/ أحمر الرجلين/ أحمر الإبل/ أحمر طيب/ أحمر أخفاها/ أحمر الموت/ أحمر الماء/ أحمر الشمس/ أحمر الرأس/ أحمر الأرض/ أحمر لونه/ أحمر كميته/ أحمر قطاة/ أحمر شقائق/ أحمر حلو/ أحمر غير/ أحمر أرض/ أحمر العيش/ أحمر القدر/ أحمر العوسج/ أحمر العمر/ أمر الطير/ أحمر السواد/ أحمر الرواح/ أحمر الحسن/ أحمر الجمال/ أحمر الأديم/ أحمر اليد/ أحمر ورق.	أبيض/ أسود / أخضر / أصفر	1428	أحمر .
أخضر اللسان/ أخضر العرب/ أخضر الأرض/ أخضر الخضرة/ أخضر اليوم/ أخضر العنب/ أخضر التهذيب/ أخضر يبس/ أخضر لونه/ أخضر عيدان/ أخضر زغريب/ أخضر خضراء/ أخضر الناس/ أخضر النازح/ أخضر الماء/ أخضر يؤكل/ أخضر نخل/ أخضر ناعم/ أخضر ناضر/ أخضر مر/ أخضر مخالط/ أخضر كالليل/ أخضر غطاء/ أخضر عظام/ أخضر عرق/ أخضر ظهر/ أخضر طيب/ أخضر سوادا/ أخضر ريان/ أخضر حمرة/ أخضر بطنه/ أخضر أروق/ أخضر غير/ أخضر أعلاه/ أخضر أحوى/ أخضر النخل/ أخضر النبق/ أخضر الثيب/ أخضر المنظر/ أخضر الكلاء/ أخضر القيقظ/ أخضر الفرس/ أخضر الطيلسان/ أخضر الشمس/ أخضر	أبيض/ أسود / أحمر / أصفر/ أزرق	658	أخضر .

متلازمات منتخبة تفصح عن تعابير اصطلاحية	التلازم مع لون آخر	عدد المتلازمات	ألفاظ الحقل اللوني
السواد/ أخضر الرجل/ أخضر الربيع/ أخضر الحمض/ أخضر الحمام/ أخضر الحديد/ أخضر الجلد/ أخضر الثمام/ أخضر التين/ أخضر الإنسان/ أخضر الإناث/ أخضر الآس.			
أصفر يابس/أصفر ورس/ أصفر الوجه/ أصفر والع/ أصفر منعقد/ أصفر مجتمع/ أصفر غليظ/ أصفر عظم/ أصفر اللون/ أصفر ناصع/ أصفر قانع/ أصفر الريح/ أصفر الحديث/ أصفر التهذيب/ أصفر النيوت/ أصفر العهد/ أصفر طيب/ أصفر الصفرة/ أصفر النطن/ ورق أصفر/ نبات أصفر/ ماء أصفر/ طائر أصفر/ أصفر غلاب/ أصفر صاف/ أصفر شديد/ أصفر خلق/ أصفر بطن/ أصفر اليد/ أصفر الروس/ أصفر النبات/ أصفر العينين/ أصفر العود/ أصفر العروق/ أصفر الظهر/ أصفر الزيت/ أصفر الرجلين/ أصفر الأصل.	أبيض / أسود/ أحمر/ أخضر	458	أصفر.
أزرق العين/أزرق الناب/أزرق عظيم/ناب أزرق/ حديد أزرق/أزرق النعرة/أزرق الطير/أزرق الصوار/ أزرق الأذن/لحم أزرق/ماء أزرق/أزرق عتيد/ أزرق صاف/أزرق شديد/حقة زرقاء/أذى أزرق/أزرق الوجه/أزرق المرار/أزرق العشب/أزرق العدو/ أزرق السموت/أزرق الرمة/أزرق الأعماع/أزرق الأثير.	أخضر، أبيض، أحمر	256	أزرق.

بصورة عامة، تعلن المتلازمات اللفظية عن ميل كلمة معينة الى الحضور مع كلمات أخرى في السياق، وبشكل متكرر. وقد ينشأ التلازم اللفظي عن إنشاء تعابير اصطلاحية، مع توسيع دائرة المعاني والدلالات من خلال إنشاء تعابير جديدة تعزز الذخيرة اللغوية. لقد نجحت الخوارزميات الذكية في النقاط مجموعة من المتلازمات اللفظية من دائرة حقل ألفاظ الألوان، في معجم لسان العرب، والتي تشكل ذخيرة لغوية جيدة، لتوظيف التعابير اللونية في سياقات متعدّدة، كما وستوفر للمستخدم فرصة توسيع مفرداته بالمزيد من التعابير الاصطلاحية التي شاع استخدامها في اللسان العربي، كما أنها ستشكل قاعدة للتعبير الاصطلاحية يمكن لنظم الذكاء الاصطناعي التوليدي استثمارها في إنتاج نصوص جديدة تتسم بمستوى بلاغي رصين في ميزان بلاغة الخطاب العربي.

ويمكن أن نستثمر هذه المتلازمات اللفظية واصطلاحاتها في توسيع دائرة الفضاء الدلالي⁹ لكل كلمة اشتقت من ألفاظ الألوان الستة، والتي يمكن أن تستخدم في نظم الذكاء الاصطناعي التوليدي، حيث يمكن أن تنتج نصوص جديدة من الذخيرة اللغوية لمعجم لسان العرب، أو تستخدم كمورد لغوي ودلالي يمكن أن توظف مفرداته والدلالات المتنوعة التي استوطنت فيها باختصار النصوص، وإنتاج المرادفات، وتوسيع دائرة النطاق الدلالي للتعبير اللونية المختلفة.

5. خلاصة:

بات واضحاً لدينا أن خوارزميات الذكاء المحوسب باتت تمتلك القدرة على التنقيب في محتوى المعاجم اللغوية العربية، وحرثة المفردات اللغوية التي تستوطن فيها. كما أن هذه الخوارزميات قد نجحت في ممارسة سلسلة من التحليلات اللسانية المحوسبة التي شملت التحليل النحوي، والصرفي، والدلالي لمادة المحتوى مع وجود أكثر من فرصة لتحليل تدفق المعاني وتعدد الدالات في حقول

معرفية مختلفة، وبزمن قياسي لا يمكن أن يقارن بما يمكن أن يستغرقه المتخصص في العلوم اللسانية لكي يبلغ ما نجحت هذه الخوارزميات الذكية في تحقيقه. كذلك فإن إعادة هيكلة محتوى معجم لسان العرب لابن منظور، ووسم هوية كياناته اللغوية قد أسهم في تحويله إلى مادة خصبة يمكن لأدوات الذكاء الاصطناعي المحوسب أن تنتج عنه مجموعة متنوعة من الفضاءات الدلالية الخصبة، لتلبية احتياجاتنا في سبر حقول معرفية محددة، وبمستويات تتباين في شمول المعالجة، وتوليد الاصطلاحات والتعبير التي ستحافظ على مفردات لغتنا العربية، وتتيح لنا أكثر من فرصة لإنتاج خطاب لغوي عربي رصين، حتى وإن غابت الفرصة عنا في قراءة ودراسة هذه المعاجم وما تحويه من ذخيرة لغوية خصبة.

مصادر الدراسة

أولاً: العربية:

ابن منظور، محمد بن مكرم بن علي الأنصاري، لسان العرب، (قم: أدب الحوزة، 1405 هـ).
 خليل، إبراهيم محمود، ألفاظ الألوان ودلالاتها عند العرب، دراسات، العلوم الإنسانية والاجتماعية، المجلد 33، العدد 3، 2006، الصفحات: 441-457.
 زيادة، وفاء، حقل الألوان في اللغة العربية: بين الدلالة المورفولوجية والدلالة المعجمية، مجلة كلية الآداب، جامعة القاهرة، المجلد (81)، العدد (4)، أبريل 2021، الصفحات: 301-365.
 عبید، كلود، الألوان: دورها، تصنيفها، مصادرها، رمزياتها، ودلالاتها، الطبعة الأولى، (بيروت: المؤسسة الجامعية للدراسات والنشر والتوزيع، 2013).

ثانياً: الإنجليزية:

Bird, Steven, Klien, Ewan and Edward Loper, **Natural Language Processing with Python**, (Cambridge: O'Reilly, 2009).
 Dilts, Robert, **Modeling with NLP**, (California: Meta Publications, 1998).

1. قيل أن ابن منظور ولد في المحرم من سنة 630 هـ بالقاهرة، وتوفي بطرابلس في شعبان من سنة 711 هـ.
2. ابن منظور، محمد بن مكرم بن علي الأنصاري، لسان العرب، (قم: مطبعة أدب الحوزة، 1405 هـ)، 1: 4.
3. حصلنا على النسخة الرقمية من مستودع المكتبة الشاملة على الإنترنت: [لسان العرب • الموقع الرسمي للمكتبة الشاملة \(shamela.ws\)](http://www.shamela.ws).

4 . Bird, Steven, Klien, Ewan and Edward Loper, **Natural Language Processing with Python**, (Cambridge: O'Reilly, 2009), pp.179,221,261,291,327,361,407.

5 . تعرّف المدونة بأنها عبارة عن نص أو مجموعة من النصوص المحوسبة التي تنتمي الى حقل من حقول المعرفة، انتخبت لكي تعبر عن الذخيرة اللغوية والاصطلاحية التي يراد سبر عناصرها في هذه النصوص، ووظفت أدوات الحوسبة الذكية لمعالجة وإدارة محتواها، وتحليل عناصرها نحويًا، وصرفياً ودلاليًا، ووسمت كياناتها لتمييز هويتها، وتحديد انتماءاتها.

6 . Dilts, Robert, **Modeling with NLP**, (California: Meta Publications, 1998), pp.29–32.

7 . عبيد، كلود، الألوان: دورها، تصنيفها، مصادرها، رمزيتها، ودلالاتها، الطبعة الأولى، (بيروت: المؤسسة الجامعية للدراسات والنشر والتوزيع، 2013)، ص:10.

8 . زيادة، وفاء، حقل الألوان في اللغة العربية: بين الدلالة المورفولوجية والدلالة المعجمية، مجلة كلية الآداب، جامعة القاهرة، المجلد (81)، العدد (4)، أبريل 2021، الصفحات: 302.

9 . يستخدم في دائرة الذكاء الاصطناعي التوليدي اصطلاح الفضاء الدلالي للكلمة Word Space الذي يعبر عنه بمتجه Vector متعدد الأبعاد، يفصح كل بعد من أبعاده في فضاء المعنى، عن معنى تمتلكه الكلمة في السياق الذي ترد فيه. ويسهم تعدد متجهات دلالة الكلمة الواحدة عن تعدد المعاني التي يمكن أن تحملها في سياقات متعددة، حيث تستثمر هذه الدلالات في توليد نصوص جديدة.