

معاجم الشيوخ والمدونات اللغوية العربية
دراسة على ضوء لسانيات المدونات
مع صناعة "معجم شيوخ مصطلحات الألوان
الأساسية في اللغة العربية"

إعداد

د. محمد أحمد عبد الصبور

أستاذ علم اللغة المساعد

قسم علم اللغة والدراسات السامية والشرقية

كلية دار العلوم – جامعة المنيا

معاجم الشبوع والمدونات اللغوية العربية
دراسة على ضوء لسانيات المدونات
مع صناعة "معجم شبوع مصطلحات الألوان الأساسية في اللغة العربية"

إعداد

د. محمد أحمد عبد الصبور

أستاذ علم اللغة المساعد

قسم علم اللغة والدراسات السامية والشرقية

كلية دار العلوم – جامعة المنيا

ملخص البحث:

سعت هذه الدراسة لتحقيق عددٍ من الأهداف النظرية، منها: إلقاء الضوء على أهمية معاجم الشبوع اللغوي لمفردات اللغة العربية. ومعرفة إلى أي مدى يمكن الاعتماد على المدونات العربية الحديثة في صناعة معاجم الشبوع اللغوي. واستيضاح أكثر المدونات العربية فعالية في الإحصاء الآلي للمفردات. وتحديد إشكاليات إحصاء الشبوع آلياً في المدونات العربية. واكتشاف المحتوى التعريفي الأمثل لمعجم الشبوع. بالإضافة إلى هدفين تطبيقيين، هما: قياس شبوع مصطلحات الألوان الأساسية في اللغة العربية. وإعداد معجم شبوع لمصطلحات الألوان الأساسية في اللغة العربية. ويمكن أن نقول: إن الدراسة قد وفقت في تحقيق أهدافها، وقدمت عدة نتائج عامة، ومنها: أنه لا يمكن الاعتماد كلياً على الباحث الآلي في المدونات العربية الكبرى في صناعة معاجم الشبوع اللغوي. والحصول على نتائج دقيقة يتطلب الفحص اليدوي من المختصين. وأن جودة الإحصاء الآلي ترتبط ارتباطاً قوياً بأن تكون جميع مفردات المدونة مشكولة شكلاً كلياً بدقة عالية. وأن ثمة إشكاليات متنوعة في الإحصاء الآلي لشبوع المفردات في المدونات العربية. كما نجحت الدراسة في تقديم إحصاء شبوع لمصطلحات الألوان الأساسية في اللغة العربية. وقدمت الدراسة "معجم شبوع مصطلحات الألوان الأساسية في اللغة العربية."

الكلمات المفتاحية:

معجم الشبوع – المدونة العربية – لسانيات المدونات – مصطلحات
الألوان الأساسية

Abstract:

This study sought to achieve several theoretical objectives, including shedding light on the importance of frequency dictionaries of Arabic language. And knowing the extent to which modern Arabic corpus can be relied upon in the creation of frequency dictionaries of Arabic. And clarifying the most effective Arabic corpus in automatic vocabulary counting. And identifying the problems of automatic counting of words in Arabic corpus. And discovering the definition components that fit the frequency dictionary. In addition to two practical objectives: measuring the prevalence of basic color terms in the Arabic language. And creating a " Frequency Dictionary of Basic Color Terms in Arabic". We can say that the study succeeded in achieving its objectives, and presented several general results, including It is not possible to rely entirely on the automated searcher in the major Arabic corpus in making frequency dictionaries. And obtaining accurate results requires manual examination by specialists. And that the quality of automated statistics is strongly linked to the fact that all corpus tokens are fully written with all short vowels. There are various problems in the automatic counting of vocabulary in Arabic corpus. The study also succeeded in presenting a survey of the prevalence of basic color terms in Arabic. The study presented "Frequency dictionary of basic color terms in Arabic".

Key words:

Frequency Dictionary – Arabic corpus – Corpus linguistics – Basic color terms

معاجمنا، بعد أن تطورت على مدى بضعة قرون من الجهد، أصبحت الآن قطعة أثرية لقدراتنا اللغوية الفطرية بقدر ما هي منتج يوضح رؤية حضارتنا للعالم (Amsler, 1982, p. 660).

١-معاجم الشبوع اللغوي.

معاجم الشبوع اللغوي Frequency dictionaries هي نوع من المعاجم تصنف فيه الوحدات المعجمية وفقاً لتكرار استخدامها في مدونة نصية تمثل اللغة ككل، أو لنمط معين من اللغة، أو لمؤلف فردي (Derocher, 1973, p. 11). وهي معاجم تعتمد في تأليفها على الإحصاءات والأرقام، لذا فهي تخرج من منطقة حدس المؤلف الخاص الذي وصف -دوماً- بأنه "المصدر النهائي للمعرفة المعجمية لمؤلف المعاجم" (Meyer et al., 1992, p. 957). إلى منطقة الإحصاء والانضباط الرقمي؛ حيث تعيد تقديم مفردات اللغة باستخدام مناهج ترتب المفردات وفق الاستعمال الفعلي للغة. (Başaran, 2022, p572)

ومعاجم الشبوع اللغوي لها أهميتها في الدرس اللغوي لا سيما مع التقنيات الحالية في المدونات اللغوية. فمعاجم الشبوع يمكن أن يجرى عليها دراسات في مجال قواعد اللغة، والبنى الصرفية، والدلالات النحوية، بالإضافة إلى تنوع الأشكال النحوية، وتناوبها. كما تقدم تلك المعاجم وصفاً مفصلاً لبعض الظواهر المهمة، وتوفر البيانات للإجابة على كثير من الأسئلة البحثية (Lyashevskaya, 2013, p. 2). حتى تلك المعاجم التي لا ترتب المداخل وفق الشبوع إذا وضعت شبوع الكلمات في الاعتبار عند صياغة التعريفات، فإن ذلك يضيف إلى دقة المعجم وقوته. تذكر دراسة أن أحد الأسباب التي زادت من أخطاء الطلاب كان استخدامهم لمعاجم تستخدم كلمات منخفضة التردد في لغة التعريف (Xu, 2012, p. 368).

كما زادت أهمية معاجم الشبوع اللغوي في مجال تعليم اللغات، حيث أظهرت الدراسات قوة العلاقة بين معرفة المفردات والمهارات اللغوية، سواء بالنسبة لمتحدث اللغة الأولى، أو لمتعلم اللغة الأجنبية. (Schmidt, 2010, p.5) حيث إنه ليس لجميع المفردات الأهمية ذاتها في الاستعمال اللغوي، وأنه - بسبب وظيفتها الرئيسية في التعبير عن الأفكار - من المرجح أن يواجه بعض المفردات أكثر من غيرها، وبالتالي فإن بعض المفردات تشكل نسبة تكرار أعلى من غيرها في الاستعمال اللغوي. (Ech-charfi, 2023, p.2) وبالنسبة لمتعلم اللغة الثانية، فإنه بحوالي ٤٠٠٠ كلمة سيتمكن متعلم اللغة الثانية من التعرف على أكثر من ٩٠٪ من الكلمات في محادثة المتحدث الأصلي. وإذا تعلم ألفي كلمة إضافية، فإن هذه النسبة ستزيد بحوالي ٣-٤٪

فقط. (Davies, 2005, P 115) فإذا عرف المتعلم الكلمات الأكثر شيوعاً في الاستعمال سيعطيها أولوية خاصة عند التعلم. وإذا عرفنا أن المتعلم أثناء استخدامه للمعجم اللغوي العادي سيواجه أيضاً من المفردات منخفضة الشبوع، بل سيواجه عدداً كبيراً من المداخل لألفاظ تميل في الاستعمال اللغوي إلى ندرة الظهور، وإذا بحثنا عن أي من هذه الألفاظ في المدونات لن تظهر إلا في المدونات الكبيرة جداً، بل إن بعض الألفاظ يميل للظهور مرة واحدة فقط، ويطلق على هذه المفردات اسم المفردة الفريدة hapax legomena. (Ech-Charfi, 2023, p. 10) هذا الواقع موجود في المعاجم اللغوية العادية لجميع اللغات، وفي العربية هناك مجموعة من الكلمات التي لم تُذكر إلا مرة واحدة في القرآن الكريم، جمعها بعض الدراسات مثل دراسة عبد المقصود (عبد المقصود، ٢٠٠٠م) بعنوان مفاريد الألفاظ في القرآن الكريم دراسة لغوية، وكل هذه الألفاظ تشرف بوجودها في المعاجم اللغوية العربية. لذا في ظل زخم المفردات نادرة الاستعمال في المعاجم اللغوية العادية، فإن معاجم الشبوع اللغوي تعد -بالنسبة لمتعلم اللغة- أداة لا تقدر بثمن. (Davies & Gardner, 2013, p. 1) ومعلومات شيوخ المفردات لها دور مركزي في تعلم اللغة. (Nation, 1990, p. 102) وقد شهدت بدايات القرن الحالي ما عرف بقوائم شيوخ الكلمات، وكانت متاحة لمعلمي اللغة ومتعلميها. على سبيل المثال، أنتج ثورندايك Thorndike (١٩٢١)، و(١٩٣٢)، وثورندايك ولورج Thorndike and Lorge (١٩٤٤) قوائم شيوخ الكلمات في النصوص المستخدمة في تعليم الأطفال الأمريكيين. كما وضع مايكل ويست Michael West (١٩٥٣) قائمة الخدمات العامة للكلمات الإنجليزية، وكانت موجهة بشكل خاص إلى متعلمي اللغة الإنجليزية الأجانب. (Davies & Gardner, 2013, p. 7) ثم في مرحلة لاحقة ظهرت معاجم الشبوع التي لا تكتفي بذكر المفردة وتكراراتها، بل تسجل دلالتها، وتقدم الأمثلة عليها. ومنها -على سبيل المثال- معجم الكلمات الشائعة في الإنجليزية الأمريكية المعاصرة الذي يعتمد على بيانات مدونة لغوية مكونة من ٣٨٥ مليون كلمة، وردت فيما يقرب من مئة وخمسين ألف نص. والمدونة متوازنة بالتساوي بين اللغة الإنجليزية المنطوقة (محادثة غير مكتوبة من الراديو، والبرامج التلفزيونية)، والأدب (الكتب، والقصص القصيرة، ونصوص الأفلام)، وأكثر من ١٠٠ مجلة شهيرة، وعشر صحف، و١٠٠ مجلة أكاديمية. (Davies & Gardner, 2013, p. 1)

ومعاجم الشبوع لها أنواع متعددة، فهي تشترك في جمعها للمفردات وفق شيوخها في اللغة، لكن تتنوع في معالجتها لهذه المفردات، فيمكن -على

سبيل المثال- أن ترتبها حسب التكرار، أو تصنف المفردات حسب التكرار إلى مجموعة متنوعة من الموضوعات، مثل الأسرة، والرياضة، والطعام. كذلك يمكن أن تكشف عن شبوع استعمال الكلمات الجديدة، كذلك الفروق بين اللهجات.

ولا ريب أن التوسع في استخدام الحواسيب وبرامج الحاسوب المساعدة الفعالة، وتجميع المدونات اللغوية الحديثة الضخمة مكن الباحثين من توفير إحصاء الشبوع والتكرارات بطرق أكثر تنوعاً وبنائج أكثر دقة. (Davies & Gardner, 2013, p. 7) كذلك سمحت تلك المدونات الكبيرة ببنائج أكثر تمثيلاً للواقع اللغوي. (حمزة، ٢٠٢١، ص١٩) وهذه العلاقة المفيدة بين معاجم الشبوع اللغوي والحاسوب قضت على مقولة: إن المعجم والحاسوب موضوعان لا يمكن جمعهما في المقالة نفسها، ولا حتى في الاقتراح نفسه. (Guthrie, et al, 1996, p63)

ومن المثير للاهتمام أن ثمة تبايناً على نطاق واسع بين حدس المتحدث الأصلي حول تقديرات شبوع الكلمات وإحصاءات شبوع الكلمات التي جمعها خبراء اللغة من المدونات. لذا فيما يتعلق بشبوع المفردات لا يجوز الاعتماد على الحدس الشخصي مطلقاً، فربما يؤدي لأسوأ العواقب، وقد يملأ المعلمون والمؤلفون أدمغة الطلاب بمئات الكلمات التي قد لا يحتاجونها في الاستعمال، ويهملون الكلمات دائمة التردد التي يحتاجها الطلاب في جميع أدوار حياتهم في كل يوم وفي كل ساعة. (طعيمة، ١٩٨٥، ص١٣٤) وبالتالي، فإن الترددات الناجمة عن تحليل المدونات الكبرى تقدم أفضل تقدير لشبوع الكلمة. (Alderson, 2007, p. 32) وأن مؤشرها التقريبي للتكرارات غالباً ما يكون هو المطلوب تماماً. (Halliday, 2008, p. 344). ويلفت لياشيفسكي Lyashevskaya الانتباه إلى الاستقرار النسبي لبيانات الشبوع في المدونات الكبيرة، والمدونات الصغيرة. وأن المدونات تتفق بشكل جيد في قيم متوسط احتمالية حدوث الحالات. وتكمن الاختلافات في بنية جمع النص من حيث الأنواع (Lyashevskaya, 2013, p. 7)

وصناعة معاجم الشبوع اللغوي لها أيضا صعوباتها التي تواجهها لا سيما إذا تعددت المعاني للمفردة، أو إذا كانت الوحدة المعجمية متعددة المفردات، بالإضافة لإشكاليات تجميع المفردات في حقول أو عائلات عند حساب الكلمات، وتحليلها. (Gardner, 2007, p. 76) بالإضافة إلى أن آليات معاجم الشبوع تختلف عن المعاجم التقليدية في نواح مهمة أساسها أن المعاجم التقليدية تميل إلى تقليل التعقيد المعجمي للغة من أجل تحديد كلمات اللغة على أساس معانيها، أما معاجم الشبوع فتهدف إلى حساب تكرارات

جميع الكلمات المستعملة. (Spinelli, 2022, p xv) ومن الناحية المنهجية يختلف معجم الشيوخ عن جميع أنواع المعاجم الأخرى من حيث إنه يعتمد بالضرورة وحصرياً على المدونات اللغوية (Miron, 1973, p. 31). كذلك نتوقع في معاجم الشيوخ أن ألفاظ التعريف بالمعجم تكون في مجملها ألفاظاً تتميز بالشيوخ كذلك. إذ لا بد من الانطلاق من الألفاظ المعروفة المكتسبة عند البحث عن ألفاظ جديدة لاكتسابها. (حمزة، ٢٠٢١، ص١٦)

وإجمالاً نقول إن معلومات شيوخ المفردات مهمة جداً، ولها فوائد لها للدرس اللغوي، وغيره، وأنه من المنطقي في نطاق التعليم إعطاء الأولوية لتعلم الكلمات التي من المحتمل أن نسمعها ونقرأها في أغلب الأحيان. (Gardner, 2007, p. 76)

كذلك يجب التنبيه على أنه لا ينبغي أن تؤدي دراسة معلومات الشيوخ إلى استبعاد المعرفة السياقية والنصية حول استعمال اللغة. (Davies & Gardner, 2013, p. 8) ولا يجب أن تعامل بيانات التكرار الموجودة في معاجم الشيوخ بوصفها المصدر الوحيد للمعلومات. (Gardner, 2007, p. 76)

٢- إشكالية الدراسة.

يعد انتقاء الكلمات من أهم الإشكاليات التي تواجه المؤلف. واللغة العربية غنية بالمفردات، لذا فإن المؤلفات العربية بشكل عام -والتعليمية بشكل خاص- تحتاج إلى عناية في انتقاء الكلمات لاسيما من ناحية شيوخ استعمالها في الدلالة المقصودة. لكن تأتي دراسة لاندو Landau لتوضح أن معظم مؤلفات النثر العربي ينقصها الانتقاء الدقيق لمجموعة الكلمات التي تضعها في المحتوى الذي تقدمه. (Landau, 1959, p. 4) لذا تتصح بعض الدراسات المؤلف ألا يكتفي بحدسه، وأن يلقي نظرة على قوائم الشيوخ في اللغة؛ لأنها ستعينه -إن أحسن استخدامهما- بطريقة عملية على انتقاء أفضل للمفردات. (طعيمة، ١٩٨٥، ص١٣)

وقد عرفت المعجمية العربية قوائم الكلمات الشائعة، ولعل أشهرها قائمة مكة للمفردات الشائعة. (جامعة أم القرى، ١٩٨٣)، وثمة قوائم أخرى أقدم مثل قائمة المفردات الأساسية للقراءة الابتدائية (فاخر، ١٩٥٣)، وقائمة المفردات الشائعة في اللغة العربية (رشدي، ١٩٥٥)، وقائمة الرصيد اللغوي الوظيفي للمرحلة الأولى من التعليم الابتدائي (اللجنة الدائمة للرصيد اللغوي، ١٩٧٥)، وقائمة المفردات الشائعة في اللغة العربية المعروفة بقائمة الرياض (عبد، ١٩٧٩). وقد سجل طعيمة خمسا وعشرين قائمة شيوخ للمفردات بدءاً من قائمة بريل Brill في ١٩٤٠ وعنوانها الكلمات الأساسية في الصحافة

العربية اليومية، وانتهاء بقائمة شمالان من إعداد مركز الشرق الأوسط للدراسات العربية في شمالان بلبنان وعنوانها قائمة الكلمات المختارة (طعيمة، ١٩٨٥، ص ص١٣٤-١٤٩). كل هذه القوائم تفاوتت من حيث دقتها، ومن حيث عدد المفردات، ومن حيث حجم المدونات النصية التي اعتمدت عليها، بل ومن حيث اعتمادها على مدونة نصية بالأساس. (ينظر: طعيمة، ١٩٨٥، ص -١٥٠) لكنها جميعاً تشترك في أنها لم تتح لها الاستعانة بالمدونات اللغوية الحاسوبية العربية الكبرى، تلك المدونات التي لم تر النور إلا في نهايات القرن العشرين.

أما فيما يتعلق بمعاجم الشبوع العربية، فقد أشار حمزة إلى وجود مشروع لصناعة معجم للشبوع في اللغة العربية من تنفيذ المنظمة العالمية للنبوض باللغة العربية، وهو مشروع يهدف لجمع الألف كلمة الأكثر شبوعاً في العربية، لكن هذا المعجم لم يخرج إلى النور قط. (حمزة، ٢٠٢١، ص١٩) وهناك معجم "شائع" للمفردات الشائعة لمتعلمي العربية، وهو مشروع لم يكتمل بعد، وتقتصر نسخته الحالية على المفردات المئة الأكثر شبوعاً. (الفيفي وآخرون، ٢٠١٦، ص ٤). وفي عام ٢٠١١م نشرت مؤسسة روتليدج Routledge معجماً ثنائي اللغة (العربية - الإنجليزية) تحت عنوان معجم الشبوع للغة العربية (A Frequency Dictionary of Arabic)، وهو معجم يستهدف الأجانب الذين يتعلمون اللغة العربية. وحوى خمسة آلاف كلمة هي الأكثر شبوعاً في العربية الفصحى الحديثة، بالإضافة إلى عدد من اللهجات العربية الأكثر انتشاراً. واستند إلى مدونة عربية خاصة بالمعجم - غير متاحة للدارسين- مكونة من ٣٠ مليون كلمة تضمنت مواداً مكتوبة، وأخرى منطوقة من العالم العربي بأكمله. (Buckwalter & Parkinson, 2014, p. 2)

ووفق نتائج بعض الدراسات التي ترى أنه يجب أن تستند معاجم شبوع المفردات إلى مدونة متوازنة لا تقل عن ٢٠ مليون كلمة. (Brysbaert, et al, 2011, p 1) فإن معجم روتليدج هو المعجم الوحيد - وفق تعريفه لنفسه- الذي اعتمد على مدونة نصية متوازنة مليونية المفردات. وفي ضوء ما سبق نجد إشكالية بارزة موجزها أنه على الرغم من عدم وجود مدونات لغوية حاسوبية إلا أن المكتبة العربية عرفت في فترة مبكرة من القرن الماضي عدداً من قوائم الشبوع اللغوي، والآن وفي وجود مدونات لغوية عربية حديثة، والإمكانات الحاسوبية الفعالة لا نجد في المكتبة العربية معجم شبوع لغوي عربي أحادي اللغة يستند إلى أي من هذه المدونات! وهذه الإشكالية طرحت عدداً من التساؤلات، منها:

- إلى أي مدى يمكن الاعتماد على المدونات العربية الحديثة في إنشاء معاجم الشبوع؟
 - أي المدونات اللغوية العربية الحديثة أكثر فعالية في إحصاءات شبوع المفردات؟
 - هل يمكن الاعتماد كلياً على الإحصاء الآلي الموجود في المدونات؟
 - ما إشكاليات إحصاء الشبوع عند صناعة المعجم؟
 - ما المحتوى التعريفي المناسب لمعجم الشبوع؟
- والحقيقة أن مناقشة هذه التساؤلات نظرياً دون خوض غمار التطبيق لا تؤتي أكلها المرجو. لهذا، وفي ضوء أن المكتبة المعجمية العربية -وفق علمي- تخلو من معاجم شبوع لغوية عامة أحادية اللغة. كما تخلو -أيضاً- من معاجم شبوع اصطلاحية أحادية اللغة. فقد فضلنا أن تجمع هذه الدراسة بين التنظير والتطبيق. لذا توجهنا لصناعة معجم شبوع اصطلاحية مختص، هو: معجم شبوع مصطلحات الألوان الأساسية في اللغة العربية. وستساعد تجربتنا في صناعة هذا المعجم على الإجابة على تساؤلات الدراسة من واقع التطبيق الفعلي.

٣- أهداف الدراسة.

نحاول في هذه الدراسة أن نحقق عدداً من الأهداف النظرية، والتي تتمثل فيما يلي:

- إلقاء الضوء على أهمية معاجم الشبوع اللغوية لمفردات اللغة العربية.
- معرفة إلى أي مدى يمكن الاعتماد الآلي على المدونات العربية الحديثة في صناعة معاجم الشبوع اللغوية.
- استيضاح أكثر المدونات العربية فعالية في الإحصاء الآلي.
- تحديد إشكاليات إحصاء الشبوع آلياً في المدونات العربية الكبرى.
- اكتشاف المحتوى التعريفي الأمثل لمعجم الشبوع.
- كذلك تسعى الدراسة لتحقيق هدفين تطبيين، هما:
- قياس شبوع مصطلحات الألوان الأساسية في اللغة العربية.
- إعداد معجم شبوع لمصطلحات الألوان الأساسية في اللغة العربية.

٤- منهج الدراسة.

نستخدم في هذه الدراسة لسانيات المدونات corpus linguistics ومعطياتها في تحليل المدونات العربية، ونستعين بأدوات المعجميات

Lexicography في صناعة معجم شبوع مصطلحات الألوان الأساسية في اللغة العربية، كما نستخدم اللسانيات الإحصائية Statistical Linguistics وآلياتها في تحليل الإحصاءات المرتبطة بالشبوع.

٥- المدونة النصية.

في المرحلة التأسيسية لهذه الدراسة قمنا بإجراء استطلاعي لنعرف عن أكثر المدونات العربية فعالية في الإحصاء الآلي. وأي المدونات العربية -مليونية المفردات- تعطي نتائج أكثر تنوعاً، ودقة في قياس شبوع المفردات. وقد قمنا بهذا الإجراء الاستطلاعي على المدونات العربية التالية:

- الموسوعة الشعرية الإماراتية.

<https://poetry.dctabudhabi.ae>

- المدونة اللغوية العربية العالمية لمكتبة الإسكندرية.

<https://www.bibalex.org/ica/ar/about.aspx>

- المكتبة الشاملة.

<https://shamela.ws>

- المدونة العربية لمدينة الملك عبد العزيز بالرياض.

<https://corpus.kacst.edu.sa>

وقمنا باختيار مفردة الفعل الثلاثي المجرد الصحيح: (طَعِمَ) للتطبيق في هذا الإجراء الاستطلاعي، ومن ثم تبين ما يلي:

الموسوعة الشعرية:

هي مدونة شعرية من إنتاج "دائرة الثقافة والسياحة في أبو ظبي" وتشكل "الموسوعة الشعرية" نموذجاً لجهود الدائرة ممثلة في "دار الكتب" في هذا المجال، وهي مدونة عمرها يزيد عن ٢٠ سنة، منذ العام ١٩٩٥م عندما صدر قرار إنشاء "الموسوعة الشعرية" حتى إطلاق إصدارها الأول في العام ١٩٩٨، حتى إعادة تدشينها بصورتها الجديدة في مارس ٢٠١٦. وتحوي الموسوعة الشعرية ما يزيد عن مليوني بيت شعر بالإضافة لمكتبة نثرية ضخمة. (<https://poetry.dctabudhabi.ae/#/about-us>). وعندما بحثنا في الموسوعة الشعرية عن إحصاءات الفعل الثلاثي (طَعِمَ) كانت النتائج كما يلي:

١- طَعِمَ (بالشكل الكلي): عدد التكرارات = ١٨٨٦

٢- طعم (بدون شكل كلي): عدد التكرارات = ١٨٨٦

٣- الباحث الصرفي الآلي: لا يوجد باحث صرفي آلي.

وهذه النتائج تفيد بأن الباحث الآلي للموسوعة الشعرية لا يفرق بين الكلمة المشكولة، وغير المشكولة، فحوت النتائج ما يلي:

المفردة مشكولة بالكامل: (لَيْسَ) طَعْمِي طَعْمَ الْأَرَانِبِ إِذْ قُلْدُ

*** لَصَ دَرُّ اللَّقَاحِ فِي الصَّنْبَرِ

المفردة غير مشكولة تمامًا: (وذقت من الشدائد كل طعم
*** وجبت من الفدافد كل وادي)

المفردة مشكولة جزئياً: (تَسْقِي تَرَى أَنْطُونِ طُعْمَةً رَحْمَةً
*** إِذْ كَانَ فِي الدُّنْيَا يَبْرُقُ وَيَرْحَمُ)

مدونة اللغة العربية العالمية لمكتبة الإسكندرية:

المدونة اللغوية العربية العالمية لمكتبة الإسكندرية هي مدونة من إنتاج مكتبة الإسكندرية، وهي إحدى المؤسسات الثقافية المصرية العالمية. وهي مدونة لغوية للعربية المعاصرة تحوي ١٠٠ مليون كلمة محللة صرفياً ونحوياً ودلالياً. (<https://www.bibalex.org/ica/ar/about.aspx>). وعند البحث في هذه المدونة عن تكرارات الفعل الثلاثي "طعم" كانت النتائج كما يلي:

- ١- طِعِمَ (بالشكل الكلي): عدد التكرارات = لا توجد نتائج
- ٢- طعم (بدون شكل كلي): عدد التكرارات = ١٥٨٣
- ٣- الباحث الصرفي الآلي: تم البحث عن "طعم" (بدون شكل كلي) مع الاختيارات التالية:

قسم الكلمة: فعلية.

قسم الكلمة الفرعي: فعل ماض.

الوزن الصرفي: فَعِل

النتيجة: عدد التكرارات = لا يوجد نتائج

وتبرز النتائج أن المدونة اللغوية العربية العالمية لا يوجد بها شكل للكلمات. كذلك الباحث الآلي الصرفي بهذه المدونة غير دقيق، حيث لم تتفق تكراراته مع تكرارات الباحث الآلي بالكلمة.

المكتبة الشاملة:

لا تعرف المكتب الشاملة نفسها على أنها مدونة لغوية، لكن يمكن الاستفادة منها كمدونة لغوية، فهي مكتبة تشمل مئات المؤلفات التراثية، والمعاصرة، وتحديث بشكل مستمر. برعاية مكتب الدعوة وتوعية الجاليات بالروضة بالرياض. ومنهج المكتبة أنها تقوم بشكل جميع مفردات المكتبة شكلاً كلياً (<https://shamela.ws>).

وعند إحصاء تكرارات الفعل الثلاثي "طعم" وجدنا النتائج التالية:

- ١- طِعِمَ (بالشكل الكلي): التكرار = ١١٣٩٩
- ٢- طعم (بدون شكل كلي): التكرار = ١١٣٩٩
- ٣- الباحث الصرفي الآلي: لا يوجد باحث صرفي آلي

وتفيد هذه النتائج أن الباحث الآلي للمكتبة الشاملة لا يفرق بين الكلمة المشكولة، والكلمة غير المشكولة، وبهذا تصبح ميزة الشكل الكلي للكلمات فاقدة لقيمتها عند الإحصاء الآلي لتكرار المفردات.

المدونة العربية لمدينة الملك عبد العزيز:

هي إحدى المشاريع الاستراتيجية لمبادرة الملك عبدالله للمحتوى العربي. يهدف المشروع إلى بناء مدونة لغوية عربية تحوي مليار كلمة مما دون بالعربية ابتداءً من العصر الجاهلي وحتى العصر الحديث، ومن مختلف المناطق والبلدان. مع الأخذ في الحسبان طبيعة وحجم النشاط الفكري لكل فترة، وتنوع أوعية النشر فيها (مخطوطات، وصحف، وكتب، ومجلات، ودوريات علمية، ...) والسائد من المجالات العلمية والفكرية المختلفة (المعتقدات، وعلوم العربية، والعلوم الطبيعية، والأدب، ...).
(<https://corpus.kacst.edu.sa/index.jsp>).

وكانت نتائج البحث عن تكرارات الفعل الثلاثي (طَعِمَ) في المدونة العربية كما يلي:

- ١- طَعِمَ (بالشكل الكلي): عدد التكرارات = ٧٩
- ٢- طعم (بدون شكل كلي): عدد التكرارات = ١١٥٠٩
- ٣- الباحث الصرفي الآلي: تم البحث عن "طعم" بدون شكل، مع اختيار البحث بجذع الكلمة، والجذع المدخل "فعل": عدد التكرارات = ١١٥٠٩

وتبين من النتائج أن المدونة العربية لمدينة الملك عبد العزيز دقيقة في البحث عن الكلمات بالشكل الكلي، حيث تمت مراجعة كل نصوص نتائج البحث بالشكل الكلي، وتبين عدم وجود أخطاء بالشكل. ولكن هذه الدقة العالية لم تساعدنا كثيرًا؛ حيث تبرز النتائج -كذلك- أن أغلب نصوص المدونة العربية غير مشكولة. كذلك تبين أن نتائج الباحث الصرفي الآلي دقيقة، وتتفق مع نتائج البحث بالكلمة.

وبناء على هذا الإجراء الاستطلاعي تبين أن المدونة العربية لمدينة الملك عبد العزيز هي المدونة الأكثر فعالية ودقة في إحصاء تكرارات المفردة العربية.

وبعد هذا الإجراء الاستطلاعي وقع الاختيار على المدونة العربية لمدينة الملك عبد العزيز لتكون المدونة النصية لهذه الدراسة، بالإضافة لعدة أسباب مهمة أخرى منها:

- أنها أكبر مدونة عربية حصيلة لغوية، حيث تجاوز عدد مفرداتها الكلي المليار مفردة.



شكل 1: إحصاءات المدونة العربية لمدنية الملك عبدالعزيز للعلوم والتقنية - الرياض

- أنها أكثر المدونات العربية تنوعاً نصياً.
- أنها أكثر المدونات العربية اتساعاً تاريخياً.
- أنها أكثر المدونات العربية امتداداً جغرافياً.

طرق البحث الآلي في المدونة العربية:

تتميز المدونة العربية بأنها المدونة الوحيدة التي تعطي نتائج دقيقة - إلى حد كبير- عند البحث بالكلمة المضبوطة بالشكل. كما تتيح هذه المدونة البحث الآلي بنوع الكلمة (اسم، فعل). كذلك تتيح المدونة البحث الآلي بجذع الكلمة. وموقع المدونة العربية يوفر طرق البحث الآلي عن طريق عدة أيقونات، كما يلي:

● أيقونة البحث. توفر هذه الأيقونة نوعين من البحث:

- 1- البحث في نصوص المدونة بكلمة واحدة أو عدة كلمات بطريقة مشابهة لمحركات البحث، حيث تظهر السياقات التي وردت فيها هذه الكلمات مع معلومات تحوي عنوان النص، والوعاء، والمجال، والموضوع، والفترة الزمنية، والبلد الذي طبع فيه النص. ويمكن للمستخدم تحديد البحث في جزء، أو عدة أجزاء محددة من المدونة من خلال المحددات الموجودة في أعلى الصفحة. كما يمكنه تحديد أكثر من عنصر من عناصر المحدد الواحد بالضغط على زر CTRL من لوحة المفاتيح، واختيار العناصر المطلوبة. ويمكن تطبيق هذه الطريقة على جميع محددات البحث.
- 2- البحث عن عناوين النصوص بواسطة كلمة أو مجموعة كلمات. وهي مشابهة للبحث في النصوص لكن تبحث في عناوين النصوص فقط حيث تظهر النصوص التي وردت في عناوينها

هذه الكلمات مع معلومات تحوي النص، والوعاء، والمجال، والموضوع، والفترة الزمنية، والبلد الذي نشر فيه النص.

● **أيقونة البحث المخصص.** تمكن هذه الأيقونة المستخدم من البحث عن الكلمات بواسطة:

١- **الجدع:** بحيث تُظهر النتائج الجذع مع صورته المختلفة بعد إضافة السوابق واللواحق الممكنة -مثل أَل التعريف والواو والباء والضمائر المتصلة- وذلك بحسب نوع الجذع - اسم أو فعل- مع معلومات عن تكرار هذه الصور في المدونة. ويمكن من خلال النقر على الكلمة الانتقال إلى الكشاف السياقي للكلمة.

٢- **رموز البديل:** البحث بواسطة رموز البديل (%) و() حيث (%) تعني أي عدد من الأحرف و() تعني حرفاً واحداً فقط. ويمكن أن يوضع أي رمز من هذين الرمزين في أي مكان من سلسلة الأحرف التي يراد البحث عنها. فعند إدخال السلسلة (م%) فهذا يعني البحث عن أي كلمة تحتوي على الحرف م. والسلسلة (م%) تعني أي كلمة أولها حرف الميم. والسلسلة (م_) تعني البحث عن أي كلمة مكونة من ثلاثة حروف أو سطرها حرف الميم وهكذا.

● **أيقونة توزيع التكرار.** يمكن من خلال هذه الأيقونة معرفة التوزيع الإحصائي لكلمة واحدة، أو كلمتين متتابعتين أو ثلاث حسب الفترات الزمنية أو حسب أوعية المدونة كما يختار المستخدم. حيث تظهر الفترة الزمنية، أو الوعاء، ويقابلها تكرار الكلمة، وتكرارها النسبي، وعدد النصوص التي وردت فيها.

● **أيقونة الكشاف السياقي.** تستخدم هذه الأيقونة لعرض السياقات التي تظهر فيها الكلمة، أو أي مجموعة من الكلمات المتتابعة ضمن مدى محدد من الكلمات السابقة واللاحقة. يتراوح عدد الكلمات السابقة أو اللاحقة من صفر، وحتى ١٥ كلمة. وتفيد هذه الأيقونة بشكل رئيسي في الكشف عن دلالات الكلمة المركزية والفرعية، وطبيعة السياقات التي تظهر فيها هذه الكلمة.

● **أيقونة التصاحب اللفظي.** تستخدم هذه الأيقونة لقياس مدى الارتباط بين الكلمة المبحوث عنها والكلمات التي تظهر معها في السياق حسب طول السياق الذي يحدده المستخدم من خلال الكلمات السابقة واللاحقة. ويمكن

للمستخدم اختيار طريقة إحصائية، أو أكثر من ثماني طرق ممكنة في النظام لقياس قوة الترابط. ويظهر النظام المعلومات التالية للمستخدم: الكلمة وتكرارها في المدونة وتكرارها في السياق، وقيمة معامل الترابط الإحصائي. ويمكن للمستخدم تصفية النتائج، والاقتصار على كلمات معينة بحسب تكرار هذه الكلمات في كامل المدونة و/أو السياق. كما يمكن أيضًا تصفية النتائج والاقتصار على الكلمات التي تحقق قيمة معينة على الأقل حسب نتائج المقاييس الإحصائية. ويمكنه ترتيب النتائج تصاعديًا أو تنازليًا بالنقر على رأس العمود الذي يريد ترتيب النتائج بناء عليه.

شكل ٢: قائمة البحث في المدونة العربية

6- مفاهيم تأسيسية

١/٦- المدونات اللغوية ولسانيات المدونات.

المدونة اللغوية language corpus، وتجمع على المدونات اللغوية language corpora يعرفها كريستال Crystal بأنها مجموعة من البيانات اللغوية، سواء كانت نصوصًا مكتوبة، أو نسخًا من الكلام المسجل، والتي يمكن استخدامها كنقطة انطلاق للوصف اللغوي، أو كوسيلة للتحقق من الفرضيات حول لغة. (Crystal & Crystal, 2008, p.117) وهذا التعريف يجعل أي مجموعة من النصوص المكتوبة، أو الشفوية التي يمكن أن تكون بمثابة نقطة مرجعية للتحليل اللغوي يمكن تصنيفها مدونة، بغض النظر عن حجمها، أو طريقة تخزينها. فعلى سبيل المثال يمكن اعتبار

مجموعة القصائد التي استخدمها النحاة العرب لوصف قواعد اللغة العربية الفصحى مدونة. (Ech-Charfi, 2023, p. 8)
أما المدونات اللغوية الحاسوبية الحديثة، فتتميز بأنها يمكن قراءتها آلياً بعدما أتاحت تكنولوجيا الحاسوب البحث في مدونات اللغة بطريقة أسرع وأسهل بكثير، مقارنة بالمدونات غير القابلة للقراءة الحاسوبية. (Ech-Charfi, 2023, p. 8) والمدونات اللغوية الحديثة تقوم على أساسين هما: مجموعة كبيرة من النصوص التي يمكن قراءتها آلياً، ونموذج معالجة إحصائي سهل التنفيذ. وكلا الأساسين يكمل كل منهما الآخر، ولا غنى عنه. (Wang et al., 2023, p. 3).

في اللغة الإنجليزية، تم جمع أول مدونة لغوية حاسوبية (مدونة براون Brown Corpus) في أوائل الستينيات. لكن كانت المدونات الحاسوبية حينها أصغر حجماً بكثير مما هي عليه الآن، على سبيل المثال احتوت مدونة براون على مليون كلمة فقط. إلى جانب ذلك لم تكن أدوات البحث فيها متطورة كثيراً، وهذا يفسر سبب تأخر نشر البحث الأول المستند إلى مدونات براون بعد سنوات من تجميعها، وهي دراسة كوتشيرا وفرانسيس (1967). (Kučera & Francis, 1967, pp. 1-464).

وقد بدأ جمع المدونات اللغوية الحاسوبية الكبيرة في أوائل التسعينيات فقط. فظهرت - على سبيل المثال- المدونة الوطنية البريطانية (BNC) المكونة من 100 مليون كلمة من النصوص المكتوبة والمنطوقة، تمثل أنواعاً مختلفة من اللغة الإنجليزية البريطانية في القرن العشرين. كذلك مدونة اللغة الإنجليزية الأمريكية المعاصرة (COCA) التي يواصل القائمون عليها إضافة نصوص جديدة سنوياً، وتتجاوز مليار كلمة تمثل أنواعاً مختلفة من اللغة الإنجليزية الأمريكية الحديثة. (Ech-Charfi, 2023, pp. 8-9).

ونشير إلى أن بعض المدونات اللغوية الحاسوبية تكون مدونات صوتية تعتمد على اللغة المنطوقة فقط، ومن ذلك المدونة البريطانية للغة الإنجليزية المنطوقة The Corpus of British Isles Spoken English (CoBISE)، وهي مدونة تحوي 112 مليون رمز من 38680 نصاً محدد الوقت للكلمات المأخوذة من 495 قناة على YouTube للمجالس المحلية أو المنظمات الحكومية المحلية الأخرى في 453 موقعاً في إنجلترا واسكتلندا وويلز وأيرلندا الشمالية وجمهورية أيرلندا. (Coats, n.d., p. 6).

كانت الإنجليزية هي اللغة الوحيدة للمدونات الحاسوبية الكبيرة في بداياتها، ثم ظهرت المدونة الوطنية للغة الروسية، المتوفرة على موقع ruscorpora.ru. (Kopotev, 2023, p. 92)، كما أن لموضوع الشبوع اللغوي شعبية خاصة في اللغويات الروسية. (Lyashevskaya, 2013, p. 92).

3). ثم توالى المدونات اللغوية الحاسوبية الكبيرة في عدد من لغات العالم (Assunção, 2019, pp. 39-57). وهناك عدد من المدونات العربية، وبعض هذه المدونات ذا حجم كبير، كما هو الحال بالنسبة للمدونة العربية لمدينة الملك عبدالعزيز للعلوم والتقنية، والتي تتكون من مليار كلمة من النصوص العربية الكلاسيكية والحديثة. (Ech-Charfi, 2023, p. 9) والمدونة اللغوية الحاسوبية لكي تقوم بوظيفتها اللسانية لا بد أن تكون مدونة معروفة الحجم، وجميع نصوصها يمكن قراءته آلياً، وعيانتها تتسم بالتوازن، بمعنى أن تكون ممثلة إلى أقصى حد لتنوع اللغة قيد الدراسة. (Kopotev, 2023, p. 35) ومن ثم فإن السمات الرئيسية للمدونة اللغوية الحاسوبية تتمثل في أنها: يمكن قراءتها آلياً، وذات حجم معروف، وتتسم بالتوازن، وتمثل اللغة. لكن من الناحية العملية، يعد إنشاء مدونة متوازنة تمثل اللغة عملية صعبة محفوفة بالمخاطر، وتمت مناقشتها على نطاق واسع في الأدبيات العلمية، حيث لا تحقق كل مدونة لغوية الدرجة المطلوبة لتمثيل اللغة. (Kopotev, 2023, p. 91) وفي هذا السياق يقدم فيلمور Fillmore ملاحظتين مهمتين: الملاحظة الأولى: لا توجد مدونة -مهما كانت كبيرة- تحتوي على معلومات حول جميع مجالات اللغة وقواعدها التي نريد استكشافها؛ وهذا ينسحب على كل ما اطلع عليه فيلمور من مدونات. الملاحظة الثانية: أن كل مدونة -مهما كانت صغيرة- أتاحت لفيلمور الفرصة لفحصها أعطته حقائق لم يكن يتخيل اكتشافها بأي طريقة أخرى. (Fillmore, 2011, p3)

وعلى الرغم من أن المدونات اللغوية الحاسوبية لا يقتصر استخدامها على اللسانيات فقط، ولها استعمالات كثيرة مثل استعمالها في البرمجة اللغوية العصبية، وفي تحليل المشاعر. (Wang et al., 2023, p. 16) إلا أن المدونات الحاسوبية الكبيرة، مثلت قفزة انطلاقاً للسانيات الحاسوبية التي يمكن تعريفها بإيجاز بأنها "استخدام التطبيقات الحاسوبية لدراسة اللغة دراسات متنوعة متعددة التخصصات". (Amsler, 1982, p. 657). فاللسانيات الحاسوبية لم تكن لتنتج دون المصادر التي يمكن قراءتها آلياً مثل المدونات اللغوية الحاسوبية. (Kuta et al., 2009, p. 319). كذلك لم تكن معالجة اللغة الطبيعية (Natural Language Processing) لتنتج بوصفها مجال دراسة بينية تتقاطع فيها الحاسوبيات مع اللسانيات. (Barkovska, 2023, p. 14).

ومجال اللسانيات الحاسوبية مجال متسع، ومرهق. فبناء أنظمة الحاسوب التي تعالج اللغة الطبيعية يتطلب حوسبة كمية كبيرة مع معلومات نحوية ودلالية شاملة ودقيقة، كما أن هناك كثيرًا من الأنظمة التي تتطلب

الكثير من المعلومات حول عدد كبير جدًا من الكلمات. (Ingria, 1986, p1) كما أن اللسانيات الحاسوبية حتى تصل إلى مرونة التفسير البشري للغة (من خلال ٦٠٠٠٠ تمثيل رمزي فقط -متوسط المفردات لخريج جامعي نموذجي- يمكن للبشر التفوق على أجهزة الحاسوب في الفهم الفوري للغة) تحتاج إلى أساليب حسابية معقدة ومرنة -أيضًا- في بيئة البرنامج. (Barkovska, 2023, p. 15).

وقد نتج عن اتساع مجال اللسانيات الحاسوبية استقلال بعض فروعها لتصبح علومًا مستقلة. كل علم منها يركز على مجال بعينه. ومن الفروع المهمة التي استقلت عن اللسانيات الحاسوبية لسانيات المدونات corpus linguistics التي بدأت بوصفها فرعًا من اللسانيات الحاسوبية، ثم نمت، حتى استقلت لتصبح تخصصًا قائمًا بذاته. وهي منهج تجريبي في طبيعته يركز على دراسة المدونات اللغوية وتحليلها وتطويرها؛ حيث تقوم بتجميع المزيد من المدونات اللغوية في لغات مختلفة، وتطوير المزيد من البرامج. (Ech-Charfi, 2023, p. 8)

ومصطلح "لسانيات المدونات" من الناحية العملية يستخدم في معنيين: المعنى الأول: هي منهج مخصص لإنشاء الهياكل، والأدوات اللازمة لمعالجة المدونات. والمعنى الثاني: هي المنهج الذي يستخدم أساليب تحليل المدونات والمدونات ذاتها لإجراء البحوث. وبطبيعة الحال يمكن لعالم لغوي معين أن يمثل كلا الاتجاهين الأول والثاني، على الرغم من أنه من الواضح أن هناك عددًا أكبر ممن يشاركون في الاتجاه الثاني. (Kopotev, 2023, p. 91)

وإلى جانب تجميع المدونات اللغوية الحاسوبية الكبيرة تعمل لسانيات المدونات أيضًا على تطوير أدوات للبحث في تلك المدونات. ومن بين أكثر هذه الأدوات فائدة أداة المطابق concordance التي تستخرج الكلمات الرئيسية في سياقها. يُمكن المطابق الباحثين من دراسة عبارات اللغة عن طريق التحقيق في قيود التواجد المشترك بين العناصر المعجمية المختلفة (مثل التلازم اللفظي collocations)، وبين العناصر المعجمية والتركيبات النحوية (أي التراكم التجميعية collostructions). يمكن إجراء الحسابات على أساس البيانات المقدمة من قبل المطابق لتحديد قوة الارتباط بين الخصائص المعجمية والصرفية والنحوية المختلفة وغيرها من السمات اللغوية. (Ech-Charfi, 2023, p. 9) ومن الأدوات قريبة الصلة بموضوع هذه الدراسة الأدوات المتعلقة بإحصاء تكرار الكلمات. وفي هذا الصدد يعد lemmatizer واحدًا من أكثر هذه الأدوات فائدة، حيث إنه يصنف جميع الرموز المميزة للأشكال المنعكسة لكلمة رئيسية في جذع lemma واحد،

ويحسب ترددها في مجموعة. بمساعدة مثل هذه البرامج، يمكن إنشاء قائمة كلمات من مجموعة يتم فيها تصنيف الجذوع (lemmas) (أو عائلات الكلمات (word families) وفقاً لتكرار حدوثها. (Ech-Charfi, 2023, p. 9) إلى جانب ذلك، ثمة أدوات تصنف الكلمات في الجزء الخاص بها من الكلام، وتصنف المميزات النحوية، والسمات الدلالية الأخرى. وكل هذا يُمكن الباحثين من تحديد قدر كبير من الانتظام في اللغة التي لن تكون ملحوظة بالحدس وحده. (Ech-Charfi, 2023, p. 8)

وفي إطار ما تقدمه لسانيات المدونات يمكن أن نميز ثلاثة نهج لسانية شائعة لتحليل المدونات، وهي:

-نهج تحليل يستنير بالمدونات: يستخدم هذا النهج البيانات التي تم جمعها في المدونة مصدرًا لأمثلة اللغة الطبيعية. بهذا المعنى، لا يختلف هذا النهج كثيرًا عن جمع الأمثلة كأساس للبحث اللغوي. ولا يتوقع التحليل الكمي العميق مع هذا النهج. وهدف هذا النهج هو السعي نحو مزيد من الموضوعية في التحليل اللغوي بتحليل مدونة لغوية ليتم تقليل دور الحدس اللغوي للباحث كما يسعى إلى إفساح المجال أمام تحليل الحقائق اللغوية بكل تنوعاتها. (Kopotev, 2023, p. 92)

-نهج تحليل قائم على المحتوى: وخصوصية هذا النهج من تحليل المدونات هو أن بيانات المدونة لا تدرس فقط من الناحية النوعية، ولكن أيضًا من الناحية الكمية. وعادةً ما يعني هذا النهج باختبار بعض الفرضيات الأولية، والتي يتم تحديد أحكامها النظرية مسبقًا، ولا تتغير أثناء الدراسة. وفي إطار هذا النهج، تصبح البيانات الكمية والتحليل الإحصائي أداة لا غنى عنها للدراسة. وهذا النهج هو الأكثر شيوعًا، بل يمكننا القول: إن غالبية الدراسات الحديثة المعتمدة على المدونات تلتزم بهذا النهج التحليلي. (Kopotev, 2023, p. 93)

-نهج تحليل يعالج قاعدة البيانات: هذا النهج يتضمن استبعاد أو الحد من الفروض النظرية المحددة سلفًا. وفي هذا النهج تكون مهمة الباحث هي إنشاء خوارزمية لمعالجة البيانات اللغوية، تكون نتيجتها -على سبيل المثال- تحليل الظواهر اللغوية، وتصنيفها دون مشاركة بشرية. ويفسر الباحث النتائج التي تم الحصول عليها، ومن ثم تصحيح الخوارزمية. ويختلف هذا النهج اختلافًا جذريًا عن النهجين السابقين من حيث إنه يتوجه نحو إلقاء نظرة جديدة على المادة اللغوية، لملاحظة الأنماط التي لم تجذب انتباه اللغويين في السابق. وفي هذا النهج تعتبر لسانيات المدونات جزءًا من تحول كبير نحو التحليل الحاسوبي للبيانات الضخمة. (Kopotev, 2023, p. 96)

معاجم الشبوع والمدونات اللغوية العربية
دراسة على ضوء لسانيات المدونات

وقد قدمت لسانيات المدونات نتائج مهمة على مستوى النظرية اللغوية العامة، ومن أمثلة هذه النتائج أنه على الرغم من أن الكلمات الشائعة عالية التردد محصورة العدد بشكل عام، إلا أن هذه الكلمات تدخل في تشكيل نسبة عالية من النصوص. وعادة ما تكون هذه الكلمات عناصر من فئة مغلقة تحمل معنى نحوياً، وبالتالي فهي لا غنى عنها لأي جملة جيدة التكوين في اللغة. ويبدو أن هذه الحقيقة تنطبق على جميع اللغات، كما يتضح من الكلمات الثمانية الأكثر شيوعاً باللغات الإنجليزية والفرنسية والعربية في الجدول التالي. (Ech-Charfi, 2023, p. 10):

	Kilgariff (2006)		Baudot (1992)		Backwalter and Parkinson (2011)	
1	the	6,187,267	de	68,373	ال	5,004,793
2	be	4,239,632	le	42,419	و	1,110,144
3	of	3,039,444	être	26,897	في	924,823
4	and	2,687,862	un	26,613	من	745,190
5	a	2,186,369	avoir	23,570	ل	584,786
6	in	1,942,315	à	23,475	ب	553,234
7	to	1,620,850	et	23,325	على	518,692
8	have	1,375,636	les	19,230	أن	303,942

شكل ٣: الكلمات الأكثر شيوعاً في العربية والإنجليزية والفرنسية (Ech-Charfi, 2023, p. 10)

في اللغات الثلاث كانت جميع الكلمات الأكثر شيوعاً من الأدوات التركيبية: أدوات التعريف، وحروف العطف، وحروف الجر. ومن الواضح أن تكرارها العالي يرجع إلى حقيقة أن الجمل في اللغة بشكل عام لا يمكن صياغتها جيداً دون استخدام الأدوات التركيبية. (Ech-Charfi, 2023, pp. 10-11)

ونشير إلى أنه بالنسبة للغة العربية فإن المدونات ولسانيات المدونات مجال خصب فسيح متسع لعشرات الدراسات والتحقيقات؛ حيث إنه بالمقارنة بلسانيات المدونات الإنجليزية، فإن لسانيات المدونات العربية لا تزال في مهدها. (McEnery et al., 2018, p. 1)

٢/٦ - مصطلحات الألوان الأساسية.

الألوان في العربية جمع لكلمة "لون" الذي عادة تشرحه المعاجم العربية بدلالاته على هيئة الشيء كالسواد والحُمْرة. (ابن منظور، ١٩٩٣، ج١٣، ص٣٩٣) وتكمن صعوبة شرح اللون بمعاجم اللغة في كونه صفة الشيء -من بياض وسواد وغيرهما- لا تُدْرَك إلا بالنظر. (عمر، ٢٠٠٨، ج٣، ص٢٠٥١) كما أن الحدود ما بين الألوان المختلفة حدود مائعة fuzzy؛ وليس هناك حد فاصل على الطيف الضوئي يحدد أين ينتهي الأحمر، وأين يبدأ البرتقالي مثلاً. (Rosch. 1975. P. 316). كما تؤكد

دراسة لطبيعة الشفرات الذهنية لفئات اللون أن كل لفظ لوني يتم تمثيله سيكولوجيًا بألوان ارتكازية focal colors، وأن بعض الدرجات اللونية للأحمر -على سبيل المثال- تبدو عند التجربة أكثر حُمرَةً من غيرها، بل أن ما ينطبق عليه لفظ اللون لا يتطابق بالضرورة مع المستوى التعريفي للون المعين. (Rosch. 1975. P. 319)

أما عن مفهوم مصطلحات الألوان الأساسية Basic Color Terms فيرتبط بعام ١٩٦٩م حيث ألف برنت بيرلن Brent Berlin، وبول كي Paul Kay كتابًا له أهميته في التعريف بالألوان بعنوان مصطلحات الألوان الأساسية. (Berlin, et al, 1969) وتبرز أهمية هذا الكتاب في أنه غير وجه النظر التي كانت ترى أن الألوان من ناحية الألفاظ والمفاهيم تخضع للنسبية الثقافية المطلقة. (Dedrick, 2013, p. 951) حيث لفت هذا الكتاب الانتباه إلى أن الخصوصيات الثقافية لا تمنع أن ثمة خصوصية لبعض الألوان. هذه الألوان ذات الخصوصية يمكن وصفها بأنها ألوان أساسية. وقد قام المؤلفان بحصر هذه الألوان في أحد عشر لونًا: (الأسود، والأبيض، والرمادي، والأحمر، والأصفر، والأخضر، والأزرق، والبرتقالي، والبنفسجي، والوردي، والبنّي). (Berlin, 1969, p. 43). إذن اللغات قد تحتوي على عشرات المصطلحات الدالة على الألوان، لكنها تحتوي فقط على عدد محدود من مصطلحات الألوان الأساسية.

ومصطلح اللون الأساسي ينطبق على اسم اللون الذي تتوافر فيه هذه الشروط: الشرط الأول: أنه اسم ينسحب على أكبر قدر من العناصر (أي ليس مثل مصطلح أشقر blonde الذي يرتبط بالشعر)، والشرط الثاني أن يتكون من وحدة معجمية واحدة monolexic (أي ليس مثل مصطلح أزرق فاتح light blue المكون من وحدتين)، والشرط الثالث أن يكون مستخدمًا بشكل مستقر، ولا لبس فيه من فئة كبيرة من المتحدثين الأصليين للغة (أي ليس مثل مصطلح فستقي فاتح chartreuse). (Hardin, 2013, p. 1)

كذلك مصطلحات الألوان الأساسية لا يشترط وجودها كلها في جميع اللغات (Berlin, 1969, p. 43). فهناك بعض اللغات بها مصطلحات لكل الألوان الأساسية. فالإنجليزية -على سبيل المثال- بها أحد عشر لونًا أساسيًا هي الأحمر red، والأصفر yellow، والأخضر green، والأزرق blue، والأسود black، والأبيض white، والرمادي gray، والبرتقالي orange، والبنّي brown، والوردي pink، والبنفسجي purple (Hardin, 2013, p. 2). وبعض اللغات -مثل

اللغة البابوانية Papuan⁽¹⁾ - لا تمتلك إلا مصطلحين للونين منها فقط.
(Levinson, 2001, p18)

كذلك في بعض اللغات قد تحدث بعض التغيرات في مصطلحات الألوان الأساسية، فعلى سبيل المثال في عام ١٩٨٤م تقرر دراسة أن لون (التركواز turquoise) في سبيله لأن يكون لونا أساسياً في اللغة الألمانية. (hardin, 1998, p. 31) وفي عام ١٩٨٨م قرر الباحثون أن اللغة الروسية بها لون أساسي أزرق إضافي وهو اللون الأزرق الكحلي (sinyi)، ويطلق أساساً على زرقة الأسطول، بالإضافة إلى الأزرق السماوي (goluboi)، ويطلق أساساً على زرقة السماء. (hardin, 1998, p. 31) وفي العربية ثمة اعتناء بالألوان وضروبها لدى عدد من علماء العربية، منهم -على سبيل المثال- أبو منصور الثعالبي الذي جعل الباب الثالث عشر من كتابه فقه اللغة وسر العربية في ضروب من الألوان. وفيه تسعة وعشرون فصلاً. (الثعالبي، ٢٠٠٢، ص ٦٨-٧٧). كذلك ثمة دراسات حديثة في اللون ودلالاته اللغوية والنفسية منها: (دياب، ١٩٨٥)، و(عمر، ١٩٧٧)، و(علي، ٢٠٠١)، و(عبد الصبور، ٢٠١٨). وفيما يتعلق بمصطلحات الألوان الأساسية قدم أحمد مختار عمر توضيحاً له وجاهته أن ثمة فرقاً بين مصطلح "الألوان الأساسية" ومصطلح "ألفاظ الألوان الأساسية"؛ فالثاني يعالج الألفاظ لا الألوان، وأساسية الألفاظ تبنى على التحليل اللغوي، أما أساسية الألوان، فتدخل فيها اعتبارات متعددة. ونحن في اللسانيات ننظر في ألفاظ الألوان الأساسية. (عمر، ١٩٩٧، ص ٣٥). ومصطلحات الألوان الأساسية في اللغة العربية في بعض الدراسات هي تسعة ألوان: الأبيض، والأسود، والأحمر، الأخضر، والأصفر، والأزرق، والبني (الأدكن)، والرمادي (الأرمد)، والأسمر. (عمر، ١٩٩٧، ص ٣٩) وهي علي هذا يعبر عنها بأحد عشر مصطلحاً للألوان الأساسية في العربية. وثمة بعض الملاحظات على قائمة مصطلحات الألوان الأساسية في العربية منها:

- خلت هذه القائمة من مصطلحات: وردي، وأرجواني، وبنفسجي، وبرتقالي؛ لأنها مصطلحات لم ترد في لغة العرب. (عمر، ١٩٩٧، ص ٣٩) على الرغم من أنها تستعمل في العربية، وتتنطبق عليها شروط اللون الأساسي. واللافت أن بعض الدراسات الإنجليزية تؤكد

1- توجد هذه اللغة بشكل أساسي- في جزيرة غينيا الجديدة في غربي الباسيفيك، ويتحدث بها نحو أربعة ملايين نسمة. (<http://www.languagesgulper.com/eng/Papuan.html>)

أن الإنجليزية اقترضت مصطلح اللون الأرجواني والبرتقالي من اللغة اللاتينية، واللغة العربية (Hardin, 2013, p. 1).

● مصطلح أسمر لا نجده في المصطلحات الأساسية للألوان عند برلين وكاي. لكنه اعتُبر لوناً أساسياً في اللغة العربية -من وجهة نظر عمر- لأهميته الخاصة وغلبته على لون البشرة العربية -استشهد على ذلك بنص لابن منظور يقول فيه: الغالب على ألوان العرب السمرة- لذا لا بد من إدخاله في الألوان الأساسية في العربية، وعدم وجوده في نموذج برلين وكاي يعود لخلو كثير من اللغات الأوربية على لفظ يدل عليه فهي بيئات يندر فيها هذا اللون. (عمر، ١٩٩٧، ص ٣٩) لكن على الرغم من وجاهة الطرح إلا أن مصطلح اللون الأسمر يخالف الشرط الأول من الشروط الحديثة للألوان الأساسية، فهذا اسم لون يرتبط بالبشرة، مثلما يرتبط اسم اللون الأشقر بالشعر.

● وضع كلمة "أدكن" بين قوسين بعد "بني"، وكلمة "أرمد" بين قوسين بعد "رمادي" فيه إشارة إلى الترادف، لكن على الرغم من التقارب في المعنى إلا أن فحص المعاجم العربية يوضح وجود فروق دلالية دقيقة بين البني، والأدكن، كذلك بين الرمادي، والأرمد، وهذا ما سيوضحه معجم الشيوخ لمصطلحات الألوان الأساسية في هذه الدراسة.

● استُبعد اللون الزيتوني على الرغم من كونه لوناً أساسياً في عدد من تصنيفات الألوان الأساسية (Evans, 1948, p232)، وعلى الرغم من كونه مستعملاً في اللغة العربية.

وعلى الرغم من الملاحظات السابقة، فإن هذه الدراسة لا تهدف إلى إعادة تصنيف مصطلحات الألوان الأساسية في العربية، ولا تهدف إلى تمييز المصطلح الأصيل منها من المصطلح المقترض. لهذا اتساقاً مع أحد أهدافها الرئيسية، وهو قياس شيوع مصطلحات الألوان الأساسية المستخدمة بالفعل في اللغة العربية، فقد فضلت تجاوز الملاحظات التي تسقط أي لون من القائمة (تجاوز ملاحظة مخالفة شرط اللون الأساسي التي تستبعد اللون الأسمر، وتجاوز وجهة نظر الترادف التي تستبعد اللون الأدكن، واللون

الأرمد. وتجاوز وجهة نظر الأصيل والمقترض التي تستبعد البرتقالي، والبنفسجي، والوردي، والأرجواني). وإثراء للمعجم، وخدمة لمستخدمه ذهبت إلى اعتماد الحد الأقصى، وليس الحد الأدنى من مصطلحات الألوان الأساسية المستخدمة في العربية. ويمكن أن نستند في هذا على أن بعض التصنيفات التي توسعت في مصطلحات الألوان الأساسية وشملت كل هذه الألوان وزيادة على اعتبار أن هناك مصطلحات ألوان أساسية ذات مرتبة أولى، وأخرى ذات مرتبة ثانوية. (Wanscher, 1978, pp 138-139) وعليه سنعمل في معجم الشبوع اللغوي في هذه الدراسة على ستة عشر مصطلحاً، وهي ثماني مصطلحات ألوان أساسية لا خلاف عليها، وهي: أحمر، أسود، أبيض، أخضر، أزرق، أصفر، بني، رمادي. وهذه المصطلحات سنميزها بعلامة النجمة بجوار مدخل المعجم على اعتبارها مصطلحات ألوان أساسية ذات مرتبة أولى، وثمانية مصطلحات ثمة ملاحظات على كونها مصطلحات لون أساسية في اللغة العربية، وهي: أدكن، أرمد، وردي، أسمر، برتقالي، بنفسجي، أرجواني، زيتوني. ولن نميزها بنجمة على اعتبارها مصطلحات لون أساسية ذات مرتبة ثانية.

٧- إحصاءات الشبوع الأولية لمصطلحات الألوان الأساسية في المدونة العربية.

١/٧- حدود مفردة البحث الآلي في المدونة العربية.

إن التدوين الحاسوبي للغة -بشكل عام- يعتمد على التمييز بين الوحدات اللغوية بوضع مسافة فارغة. وهذا التدوين الحاسوبي جعل اللسانيين -في ظني- يحتاجون إلى تعريفات جديدة لتلك الوحدة اللغوية التي توضع بين مسافتين. فإذا كانت اللسانيات اعتنت -من قبل- بتوصيف الوحدة الدلالية الصغرى، أو أصغر وحدة لغوية ذات معنى، وفي هذا الإطار ظهر مصطلح مورفيم Morpheme والذي أفضل تعريبه بمصطلح صرفيم، فإن اللسانيات الحاسوبية وضعت الوحدة اللغوية الموضوعية بين مسافتين محل اعتناء، وهنا ظهر مصطلح توكين Token والذي أفضل تعريبه بمصطلح "الرمز". والرمز في الكتابة يعني أي وحدة لغوية موضوعية بين مسافتين بغض النظر عن طولها، أو عدد صرفيماتها. (Berg, 2014, p. 201)

ومصطلح الرمز token مناسب من ناحية التدوين لطريقة تدوين أغلب اللغات؛ حيث إن أغلب اللغات تفصل بين الكلمات بالتباعد بينها، لكن من ناحية التركيب والدلالة، فثمة اختلاف بين اللغات لما يوضع بين المسافتين الفاصلتين هل هو كلمة واحدة أو أكثر. (Ech-charfi, 2023, p.3) ودون الدخول في مقارنات بين العربية وغيرها من اللغات نقول: إن التدوين اللغوي

في العربية -من وجهة النظر الحاسوبية- يمكن أن نرى فيه -على سبيل المثال- ما يلي:

١. صرفيمات تدون مستقلة، وكل صرفيم يدون موضوعاً بين مسافتين، مثل: (إنّ، أخذ، أعطى).

٢. صرفيمات متصلة دلاليّاً وتركيبياً وإملائيّاً تدون موضوعاً بين مسافتين كأنها وحدة لغوية واحدة، مثل: (إنهم، أخذناهم، أعطيناكموه).

٣. صرفيم ظاهر دلاليّاً لكنه مضمّر مستتر إجبارياً في البنية العميقة للكلام. ومن ذلك كل فعل مضارع يبدأ بألف المتكلم أو نون المتكلمين (أقوم، نقوم). فالبنية السطحية تحوي فعلاً مضارعاً، والبنية العميقة تحوي فعلاً مضارعاً + الفاعل. وفي تفسير هذه الظاهرة يقول حسان: "إن وجود المضارعة يكون حاسماً في الدلالة على الشخص، ومن ثمّ يكون استتار الضمير واجباً، فالهمزة مثلاً تدل على المتكلم المفرد، ومن ثمّ يستتر الضمير وجوباً في "أقوم"، والنون تدل على المتكلمين، ولا يشاركتها في ذلك غيرها، ومن ثمّ يكون استتار الضمير واجباً في "نقوم" (حسان، ٢٠٠٦، ص ١٥٩).

إذن البحث الحاسوبي عن مفردة عربية لا يستطيع الاعتماد على مفهوم الرمز token لتحديد بداية الكلمة ونهايتها، فالبحث عن كلمة بعينها وفق مفهوم الرمز يمكن أن يكون دقيقاً في الحالة الأولى فقط حيث الكلمة العربية تمثل صرفيماً واحداً ظاهر، ومدون، ومستقل بين مسافتين. لكن في حالات أخرى، فإن البحث الحاسوبي الآلي لا يمكن الاطمئنان لدقته. ففي الحالة الثانية يتبين لنا صعوبة البحث الدقيق آلياً عن الضمير "نا"، والضمير "هم"، والضمير "كم"، والضمير "و"، والضمير "ه". أما الحالة الثالثة فيمكن أن نتفهم صعوبة البحث الدقيق عن الضمير "أنا" والضمير "نحن" في الفعلين "أقوم" و"نقوم"؛ فكيف نحصي الحالات التي يكون فيها الضمير مستتراً وجوباً غير ظاهر في البنية السطحية، مثبتاً دالاً له وظيفة نحوية في البنية العميقة، وكذلك الحال في كل ضمير مستتر وجوباً.

كذلك تقترح الدراسات اللسانية مصطلح lemma التي تجمع على: lemmas أو lemmata، والذي يمكن أن نترجمه بمصطلح "جذع"، ومفهوم هذا المصطلح هو "كلمة رئيسية، والصيغ المنبثقة عنها". (Francis, 1982, p22) فإذا نظرنا إلى فعل مثل "كُتِبَ" فهو lemma يمثل الصيغ المنبثقة عنه، مثل: يكتب، اكتب، كتاب، مكتوب، كاتب، كتابان، كُتِبَ، مكتبة... إلخ.

عادةً ما يكون شكل الكلمة الذي يتم اختياره ليكون بمثابة lemma هو الصيغة الأقل تميزاً. (Francis, 1982, p22) وفي اللغة العربية عادة ما يكون الثلاثي المفرد المذكر الغائب الماضي هو الصيغة الأقل تميزاً. (Francis, 1982, p22) وقد أكدت الدراسات أن الملائم للمعالجة الحاسوبية للغة العربية أن يكون الفعل الماضي المجرد هو الأصل في الاشتقاق. (علم، ٢٠٢١، ص ٢٩) وقد فضلنا ترجمة Lemma بمصطلح الجذع لأن الجذع أصل الشيء، كما أنه يتشابه مع الجذر في كونه أصل للصيغ المنبثقة عنه، لكن الجذع يعتد بالمفردة في الاستعمال الفعلي، أما الجذر فيعتد بأصل المفردة قبل أن ينالها إعلال أو إبدال. فعلى سبيل المثال "قال" جذع، أما الجذر فهو "ق و ل".

إن البحث في المدونات اللغوية الحاسوبية بالجذع يعطي نتائج تشمل الجذع، وصوره المنبثقة عنه، وما يلحق به من سوابق ولو احق. لكن ما مدى دقة هذا البحث في المدونات العربية مع المفردات التي تحوي أكثر من معني مثل خوى، التي تعني جاع، وتعني خلا، وكلمة عين التي تعني البئر أو الجاسوس أو المبصرة؟ كذلك ما مدى دقة هذا البحث مع المفردات المتشابهة بنية وضبطاً، لكنها لا تنتمي للجذع ذاته، وتختلف دلالة واشتقاقاً؟ مثل كلمة (عدا) فهي فعل ماضٍ مرادف للفعل جرى: "عدا يعدو عدواً وعدواً" (ابن دريد، ١٩٨٧، ج ٢، ص ٦٦)، و(عدا) من العدوان: عدا عليه، من العدوان، يعدو عدواً وعدواً وعدواناً، إذا جار. (ابن دريد، ١٩٨٧، ج ٢، ص ٦٦)، وهي كذلك أداة جر تدل على الاستثناء. وفي مثل هذه الحالات لا يستطيع الضبط بالشكل أن يفيد، ولا بد مع هذه الحالات، ومثيلاتها من الكشف البشري عن السياقات.

إن يتبين مما سبق أن حدود المفردة العربية لها خصوصية تجعل البحث الآلي الخالص الذي لا يعتمد على توكيد من فحص يدوي للباحث المتخصص يصعب الاطمئنان لصحة نتائجه، لا سيما في الحالات الدلالية الخاصة للمفردات اللغة العربية مثل المشترك اللفظي على سبيل المثال، أو حالات التدوين الخاصة التي يتشابه فيها تدوين المفردات، وتتباعد معانيها ومشتقاتها.

وفي ضوء ما سبق سنقوم باختيار مفردة البحث على أساس مفهوم الجذع Lemma لأنه المفهوم الأكثر مناسبة للبحث الآلي عن مصطلحات الألوان الأساسية في المدونة العربية. مع التنبيه إلى خصوصية المفردة العربية، والتي قد تتطلب منا المراجعة اليدوية لكافة سياقات كل مصطلح من المصطلحات.

٢/٧- خيارات البحث الآلي عن مصطلحات الألوان الأساسية في المدونة العربية.

في هذا الموضوع من الدراسة نبين طريقتنا في البحث عن مصطلحات الألوان الأساسية في المدونة العربية وفق خيارات البحث الآلي المناسبة المتاحة في المدونة العربية.
أولاً- الخيارات العامة:

١- البحث بالمصطلح مشکولاً، وغير مشکول:

حيث إن من مميزات المدونة العربية أنها تعتمد البحث بالمفردة المشكولة، ولا تهمل حركات الشكل عند البحث، ورغبة في دقة النتائج قمنا بادئ الأمر بالبحث عن المصطلح مشکولاً شكلاً كلياً بدون حركة الإعراب: (أبيض، أسود، ... إلخ).

وبعد استخراج جميع النتائج للمصطلحات المشكولة قمنا بإعادة البحث على أن يكون المصطلح في خانة البحث غير مشکول، وذلك لمعرفة الفروق الإحصائية -إن وجدت- بين البحث بالشكل، والبحث بغير شكل.

٢- المسح الشامل للمدونة العربية:

اتبعنا عند البحث عن شيوخ مصطلحات الألوان الأساسية أسلوب المسح الشامل للمدونة. ولم نحدد عينة مخصصة، أو فترة زمنية، أو منطقة جغرافية، أو نصوصاً بعينها. فكل مصطلح بحثنا عنه في المدونة كاملة.

ثانياً- الخيارات الخاصة:

١-خيارات البحث بالجذع.

تتيح أيقونة البحث بالجذع أن نكتب في خانة البحث جذع المصطلح (أبيض، أسود ... إلخ)، ثم يتولى الباحث الآلي البحث عن الجذع، وعن بقية الصور المنبثقة عن هذا الجذع. وقمنا عند البحث بجذع المصطلح بتحديد الاختيارات التالية:

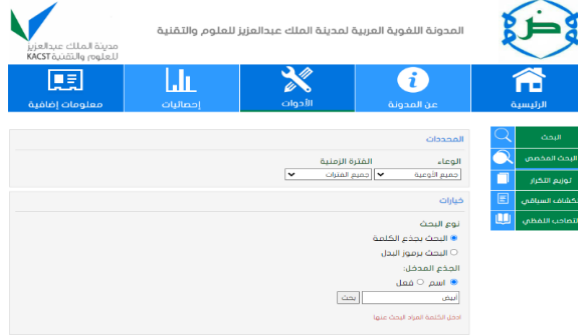
اختيار ١- جميع الأوعية.

اختيار ٢- جميع الفترات.

اختيار ٣- البحث بجذع الكلمة.

اختيار ٤- الجذع المدخل اسم.

معاجم الشبوع والمدونات اللغوية العربية دراسة على ضوء لسانيات المدونات



شكل ٤: اختيارات البحث عن جذع مصطلح اللون الأساسي في المدونة العربية

ومن ثم تظهر نتائج البحث وفق الاختيارات المحددة بهذا الشكل:

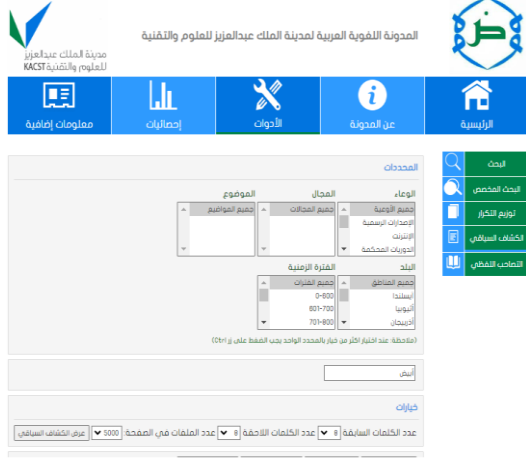
ظهر الجذع أيضا 17451 مرة في هذا المحدد			
التكرار	بالوواجق	التكرار	بالسوايق
44	الأبيضين	247	الأبيض
36	الأبيضان	139	والأبيض
33	وأبيضها	85	وأبيض
27	وأبيضه	39	بأبيض
22	وأبيضهم	29	للأبيض
10	بأبيضهم	4	بأبيض
5	بأبيضها	2	لأبيض
5	والأبيضان	1	فالأبيض
4	بأبيضه	1	كألأبيض
3	بأبيضين	1	فأبيض
3	وأبيضهما		وبالأبيض
3	لأبيضين		كأبيض
2	وأبيضك		ولأبيض
2	ولأبيضه		ألأبيض
2	لأبيضهم		أبيض
1	بأبيضهم		وكالأبيض
1	فالأبيضان		
1	لأبيضه		
1	الأبيضين		

شكل ٥: نتائج البحث بالجذع في المدونة العربية وفق اختيارات البحث

2- خيارات البحث بالكشاف السياقي:

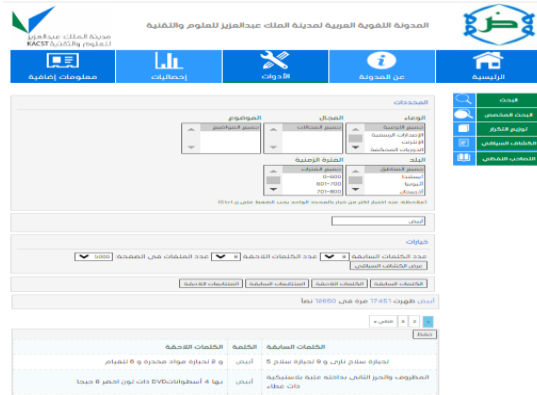
تتيح أيقونة الكشاف السياقي في المدونة العربية البحث عن المصطلح في سياقاته التي ورد فيها. ولضمان دقة نتائج التكرارات نقوم بعملية البحث السياقي لكل مصطلح من أجل مراجعة معانيه داخل سياقها، والتأكد من أنه اللفظ المقصود في معجمنا. وكانت اختيارات البحث بالكشاف السياقي كما يلي:

- اختيار ١- جميع الأوعية.
- اختيار ٢- جميع المجالات.
- اختيار ٣- جميع المواضيع.
- اختيار ٤- جميع البلدان.
- اختيار ٥- جميع الفترات الزمنية.
- اختيار ٦- عدد الكلمات السابقة على اسم اللون ثماني كلمات.
- اختيار ٧- عدد الكلمات التالية على اسم اللون ثماني كلمات.



شكل ٦: اختيارات البحث في الكشف السياقي عن الألفاظ الأساسية للألوان

ومن ثم تخرج نتائج البحث بالكشف السياقي بالشكل التالي:



شكل ٧: نموذج لنتائج الكشف السياقي في المدونة العربية

والبحث بالكشف السياقي نستعمله في الكشف عن حقيقة استعمال المفردة في النص. وبعد تدقيق جميع السياقات التي ورد فيها اللفظ يتم إعادة احتساب التكرارات في حال استبعاد أي لفظ لا يستعمل في دلالاته على اللون.

٣/٧- النتائج الأولية للقياس الآلي لشبوع مصطلحات الألوان الأساسية في
المدونة العربية:

وفقاً لمحددات مفردة البحث الآلي، ثم خيارات البحث الآلي عن
مصطلحات الألوان الأساسية في المدونة العربية أسفرت علميات البحث
الآلي عن قيم التكرارات التالية:

تكراره	أكثر صوره استعمالاً	مجموع التكرارات	صوره في الاستعمال اللغوي بالسوابق واللواحق			تكراره		اسم اللون	
			بالسوابق واللواحق	باللواحق	بالسوابق	تكرار المشكول/ غير المشكول	غير مشكول مشكول		
40188	الأبيض	63181	الأبيضين, 44 الأبيضان, 36 وأبيضها, 33 وأبيضه, 27 وأبيضهم, 22 بأبيضهم, 10 بأبيضها, 5 والأبيضان, 5 بأبيضه, 4 بالأبيضين, 3 وأبيضهما, 3 لأبيضين, 3 وأبيضك, 2 ولأبيضه, 2 لأبيضهم, 2 بأبيضهن, 1 فالأبيضان, 1 لأبيضه, 1 الأبيضين, 1 والأبيضين, 1 وأبيضكم, 1	أبيضين, 247 أبيضان, 133 أبيضه, 85 أبيضها, 39 أبيضهم, 29 أبيضة, 4 أبيضون, 2 أبيضكم, 1 أبيضتها, 1 أبيضين, 1 وبالأبيض, 26 كأبيض, 24 وللأبيض, 12 الأبيض, 11 أبيض, 2 وكالأبيض, 2	الأبيض, 40186 والأبيض, 2076 وأبيض, 1020 بالأبيض, 842 للأبيض, 245 بأبيض, 230 لأبيض, 144 فالأبيض, 76 كالأبيض, 47 فأبيض, 38 وبالأبيض, 26 كأبيض, 24 وللأبيض, 12 الأبيض, 11 أبيض, 2 وكالأبيض, 2	بوح	17451	56	أبيض

49708	الأحمر	70067	الأحمر	161, أحمران, 79, أحمرين, 71, أحمره, 51, أحمره, 50, أحمرها, 33, أحمر ي, 24, أحمر هم, 7, أحمر ون, 2, أحمر كم, 1, أحمر تك, 1, أحمر هما, 1, أحمر ته, 1, أحمر ات, 65, الأحمرين, 44, الأحمران, 39, وأحمره, 34, وأحمره, 28, وأحمرها, 13, للأحمر ي, 12, الأحمره, 11, بأحمرها, 9, وأحمر هم, 9, والأحمر ي, 8, وأحمر ي, 7, والأحمران, 6, بأحمره, 5, كأحمر ي, 3, الأحمره, 3, بالأحمرين, 3, والأحمرين, 3, لأحمره, 2, والأحمره, 2, بأحمر ي, 2, كالأحمرين, 2, كالأحمر ي, 2, بالأحمر ي, 1, كالأحمره, 1, وأحمر كم, 1, فأحمره, 1, وأحمرين, 1, وأحمر ون, 1, كأحمره, 1, بالأحمره, 1, لأحمر ي, 1, الأحمر ي, 1, بأحمرين, 1, للأحمره, 1, الأحمر ات, 1, كأحمر هم, 1, بأحمره, 1, الأحمر هم,
			الأحمر, 49708, والأحمر, 1549, وأحمر, 717, بالأحمر, 601, بأحمر, 202, للأحمر, 166, كأحمر, 136, لأحمر, 131, كالأحمر, 73, فالأحمر, 52, فأحمر, 42, الأحمر, 11, وبالأحمر, 7, وللأحمر, 6, وبأحمر, 2, فبالأحمر, 1, وكأحمر, 1,	

معاجم الشبوع والمدونات اللغوية العربية
دراسة على ضوء لسانيات المدونات

المدونة	المدونة	المدونة	المدونة	المدونة	المدونة	المدونة	المدونة		
26581	الأخضر	37568	الأخضري, 122 الأخضرين, 90 وأخضره, 14 للأخضري, 13 الأخضران, 10 وأخضرها, 8 بأخضره, 6 بأخضرها, 5 بالأخضري, 5 فأخضره, 2 فأخضرها, 2 لأخضره, 2 والأخضري, 2 وأخضرة, 2 بالأخضرين, 2 والأخضران, 1 والأخضرين, 1 الأخضرها, 1	أخضران, 206 أخضرين, 75 أخضرها, 29 أخضره, 23 أخضري, 7 أخضرها, 4 أخضرك, 4 أخضرين, 1 أخضرون, 1 أخضرتها, 1 أخضرها, 1	الأخضر, 26581 والأخضر, 1237 للأخضر, 662 وأخضر, 330 بالأخضر, 293 فالأخضر, 54 بأخضر, 51 كالأخضر, 44 لأخضر, 29 فأخضر, 19 الأخضر, 9 كأخضر, 6 وللأخضر, 5 وبالأخضر, 5 وكأخضر, 3 أأخضر, 1 أو أخضر, 1	أخضر	7598	51	أخضر
106	أدكن	177	لا يوجد	أدكنها, 1	الأدكن, 47 وأدكن, 11 والأدكن, 6 بأدكن, 3 كالأدكن, 1 كأدكن, 1 لأدكن, 1	أدكن	106	2	أدكن
401	الأرجواني	887	الأرجوانية, 176 وأرجوانية, 8 والأرجوانية, 4 الأرجوانيتين, 1 فأرجوانية, 1	أرجوانية, 116 أرجوانيتين, 1 أرجوانيات, 1 أرجوانيتان, 1	الأرجواني, 401 والأرجواني, 55 وأرجواني, 5 بالأرجواني, 2 وبالأرجواني, 2 كالأرجواني, 1	أرجواني	112	0	أرجواني

أرمد	أرمد	أرمد	أرمد	أرمد	أرمد	
377	501	1146 الأزرق، 94 للأزرق، 93 والأزرق، 21 الأزرقين، 10 وأزرقه، 7 الأزرقان، 5 وأزرقها، 4 بأزرقها، 4 وأزرق، 3 بالأزرق، 2 لأزرق، 2 بأزرقه، 2 الأزرقات، 2 وللأزرق، 2 كالأزرق، 2 الأزرق، 1 فازرقك، 1 كأزرق، 1 لأزرقين، 1 للأزرقين، 1 بالأزرقين، 1 فازرقها، 7 الأرمدة، 5 وأرمدة، 2 فأرمدها، 1 كأرمدة،	أرمدة، أرمدهما، أرمدتين، أرمدم، أرمدتها،	65، الأرمدة، 10، وأرمدة، 8، والأرمدة، 5، لأرمدة، 3، فالأرمدة، 3، كالأرمدة، 2، بأرمدة، 2، للأرمدة،	377	4
15027	21490	59 أزرقان، 18 أزرقها، 17 أزرقه، 14 أزرقين، 12 أزرق، 4 أزرق، 3 أزرقك، 2 أزرقيا، 1 أزرقات، 1 أزرقكم، 1 أزرقهما،	الأزرق، والأزرق، للأزرق، بالأزرق، وأزرق، لأزرق، بأزرق، فالأزرق، كالأزرق، فأزرق، وبالأزرق، الأزرق، كأزرق، أوأزرق، وبأزرق،	15027، 1010، 262، 220، 167، 32، 32، 28، 25، 5، 5، 5، 2، 1،	3127	4
					أزرق	

معاجم الشبوع والمدونات اللغوية العربية
دراسة على ضوء لسانيات المدونات

الأسمر	7473	الأسمر ي، 1833	أسمر ة، 421 أسمر ي، 11 أسمر ه، 8 أسمر ها، 4 أسمر ان، 2 أسمر ك، 2 أسمر ين، 1	الأسمر، 2709 وأسمر، 145 والأسمر، 69 بأسمر، 65 بالأسمر، 21 فأسمر، 18 لأسمر، 12 كالأسمر، 7 فالأسمر، 4 وبالأسمر، 3 للأسمر، 3 وبأسمر، 2 كأسمر، 1 وكالأسمر، 1	اسمر	2001	5	أسمر
2709	7473	الأسمر ي، 1833 وأسمر ة، 34 بأسمر ة، 25 لأسمر ة، 15 والأسمر ي، 15 للأسمر ي، 12 الأسمر ان، 8 الأسمر ين، 6 وأسمر ه، 3 وأسمر ها، 2 والأسمر ين، 1 فأسمر ه، 1 لأسمر ي، 1 فأسمر ة، 1 فأسمر ها، 1 والأسمر ة، 1 بأسمر ها، 1 الأسمر ي، 1 وبأسمر ي، 1 بالأسمر ي، 1	أسمر ة، 421 أسمر ي، 11 أسمر ه، 8 أسمر ها، 4 أسمر ان، 2 أسمر ك، 2 أسمر ين، 1	الأسمر، 2709 وأسمر، 145 والأسمر، 69 بأسمر، 65 بالأسمر، 21 فأسمر، 18 لأسمر، 12 كالأسمر، 7 فالأسمر، 4 وبالأسمر، 3 للأسمر، 3 وبأسمر، 2 كأسمر، 1 وكالأسمر، 1	اسمر	2001	5	أسمر
50556	80131	الأسودين، 328 الأسودان، 196 والأسود ها، 70 والأسود هم، 56 الأسود ة، 54 بالأسود ين، 41 والأسود ة، 33 والأسود ك، 26 والأسود ه، 24 للأسود ين، 18 الأسود ي، 17 والأسود ان، 13 بأسود ه، 11 بأسود ها، 10 لأسود ها، 9 والأسود ين، 6 والأسود ين، 5 لأسود ه، 3 والأسود ك، 2 بأسود ين، 2 كالأسود ان، 2 لأسود ك، 2 والأسود ة، 2 فأسود هم، 1 كالأسود ين، 1 والأسود ان، 1 الأسود ين، 1 بأسود هم، 1 والأسود هن، 1	أسودين، 226 أسودان، 163 أسود ة، 157 أسود ه، 113 أسود ها، 87 أسود هم، 23 أسود ك، 17 أسود ي، 8 أسود هم، 5 أسود كهم، 4 أسود هن، 4 أسود تهم، 4 أسود ات، 1 أسود هما، 1	الأسود، 50556 والأسود، 5013 والأسود، 2114 بالأسود، 672 للأسود، 373 بأسود، 189 كالأسود، 174 لأسود، 169 فالأسود، 60 كأسود، 57 وللأسود، 21 فأسود، 21 الأسود، 19 وبالأسود، 16 وكالأسود، 3 ولأسود، 2 وبأسود، 1 أوأسود، 1 فلالأسود، 1	اسود	18920	209	أسود

7896	الأصفر	13894	الأصفرين, 31 وأصفره, 21 الأصفران, 20 وأصفرها, 9 وأصفر ي, 8 وأصفر هم, 7 الأصفر ي, 5 والأصفرين, 4 بأصفره, 3 والأصفران, 3 بأصفرها, 3 للأصفرين, 1 وأصفران, 1 وأصفرهما, 1 كأصفر ك, 1 بأصفر كم, 1	أصفرين, 28 أصفران, 22 أصفره, 17 أصفرها, 9 أصفر ي, 8 أصفرته, 6 أصفر هم, 1	الأصفر, 7896 والأصفر, 1061 وأصفر, 472 بالأصفر, 133 للأصفر, 127 بأصفر, 41 فالأصفر, 32 كالأصفر, 15 فأصفر, 10 وللأصفر, 3 وبالأصفر, 3 الأصفر, 3 لأصفر, 2 كأصفر, 2	3884	23	أصفر
2397	البرتقالي	4335	البرتقالية, 657 والبرتقالية, 49 وبرتقالية, 41 البرتقاليين, 17 البرتقاليون, 10 البرتقاليات, 3 للبرتقاليين, 3 البرتقاليين, 3 البرتقالية, 3 فيرتقالية, 1 بالبرتقالية, 1	برتقالية, 289 برتقاليين, 4 برتقاليان, 1 برتقاليون, 1 برتقاليين, 1	البرتقالي, 2397 والبرتقالي, 374 للبرتقالي, 89 وبرتقالي, 31 بالبرتقالي, 30 فالبرتقالي, 14 كالبرتقالي, 10 فيرتقالي, 1	305	0	برتقالي
2301	البنفسجي	4228	البنفسجية, 2301 والبنفسجية, 39 وبنفسجية, 31 البنفسجيين, 7 البنفسجية, 6 بالبنفسجية, 2 البنفسجيتان, 2 بنفسجية, 2 والبنفسجيين, 1 كالبنفسجية, 1 البنفسجيون, 1	بنفسجية, 374 بنفسجيات, 2 بنفسجيتان, 1 بنفسجيتين, 1 بنفسجيين, 1	البنفسجي, 898 والبنفسجي, 217 وبنفسجي, 22 للبنفسجي, 12 بالبنفسجي, 10 فالبنفسجي, 5 كالبنفسجي, 4	288	0	بنفسجي

معاجم الشبوع والمدونات اللغوية العربية
دراسة على ضوء لسانيات المدونات

38338	البنية	290054	بنية, 31230	بنية, 14342	بنية	213117	2	ب
	البنية, 38338		بنية, 31230	بنية, 14342	بنية	213117	2	ب
	البنيان, 4781		بنية, 4497	بنية, 9177	بنية	213117	2	ب
	والبنية, 4057		بنيتها, 3465	بنية, 3615	بنية	213117	2	ب
	أبنية, 3048		بنيان, 3355	البنية, 3118	بنية	213117	2	ب
	للبنية, 2054		بنيتها, 2886	أبنية, 1263	بنية	213117	2	ب
	وبنية, 1725		بنياته, 1236	ولبنية, 560	بنية	213117	2	ب
	البنيات, 1594		بنيات, 1113	والبنية, 443	بنية	213117	2	ب
	بالبنية, 1391		بنياتها, 746	كبنية, 419	بنية	213117	2	ب
	لبنية, 1145		بنيتها, 625	فبنية, 283	بنية	213117	2	ب
	ببنية, 956		بنيك, 618	للبنية, 69	بنية	213117	2	ب
	وبنيه, 948		بنياتهم, 346	بالبنية, 47	بنية	213117	2	ب
	لبنيه, 899		بنيتهم, 274	وببنية, 30	بنية	213117	2	ب
	كالبنيان, 818		بنياته, 213	كالبنية, 22	بنية	213117	2	ب
	وبنيتها, 604		بنياتها, 205	البنية, 10	بنية	213117	2	ب
	وبنيتها, 576		بنيتهم, 172	فبنية, 8	بنية	213117	2	ب
	أبنيتها, 326		بنيتين, 116	وكبنية, 4	بنية	213117	2	ب
	أبنيتها, 302		بنيتهما, 106	قالبنية, 4	بنية	213117	2	ب
	والبنيان, 296		بنيتهما, 87	أبنية, 2	بنية	213117	2	ب
	كبنية, 283		بنياتهم, 82	فبنية, 1	بنية	213117	2	ب
	والبنيات, 262		بنيتك, 66	أو بنية, 1	بنية	213117	2	ب
	فالبنية, 195		بنيتكم, 51	أو كبنية, 1	بنية	213117	2	ب
	بالبنيان, 188		بنياتكم, 47	وبالبنية, 1	بنية	213117	2	ب
	وبنيان, 178		بنياتك, 42	فكبنية, 1	بنية	213117	2	ب
	البنيتين, 170		بنية, 41		بنية	213117	2	ب
	ببنية, 158		بنيتان, 30		بنية	213117	2	ب
	أبنيتهم, 149		بنيك, 21		بنية	213117	2	ب
	بنيتها, 130		بنياتكم, 18		بنية	213117	2	ب
	ببنيتها, 124		بنيتكم, 16		بنية	213117	2	ب
	للبنيان, 116		بنيون, 13		بنية	213117	2	ب
	البنية, 115		بنيين, 11		بنية	213117	2	ب
	وبنيات, 113		بنياتهما, 10		بنية	213117	2	ب
	وبنياتهما, 112		بنيهن, 10		بنية	213117	2	ب
	للبنيات, 103		بنيتهن, 8		بنية	213117	2	ب
	لبنيان, 99		بنياتهما, 6		بنية	213117	2	ب
	لبنيتها, 97		بنيتكما, 6		بنية	213117	2	ب
	وبنيتهم, 96		بنياتك, 5		بنية	213117	2	ب
	بنيان, 91		بنيتين, 2		بنية	213117	2	ب
	لبنيتها, 89		بنياتنا, 1		بنية	213117	2	ب
	وبنياته, 89				بنية	213117	2	ب
	لبنيتها, 88				بنية	213117	2	ب
	أبنيه, 84				بنية	213117	2	ب
	وبنها, 72				بنية	213117	2	ب
	ولبنيه, 72				بنية	213117	2	ب
	أبنيتها, 69				بنية	213117	2	ب
	فبنية, 63				بنية	213117	2	ب
	لبنيات, 62				بنية	213117	2	ب
	وبنيك, 61				بنية	213117	2	ب
	بالبنيات, 58				بنية	213117	2	ب
	لبنيتك, 57				بنية	213117	2	ب
	وبنياته, 55				بنية	213117	2	ب
	ببنيات, 52				بنية	213117	2	ب
	وبنياتهما, 44				بنية	213117	2	ب
	ببنياته, 43				بنية	213117	2	ب
	كبنيان, 38				بنية	213117	2	ب
	كالبنية, 33				بنية	213117	2	ب
	وبنيتهم, 32				بنية	213117	2	ب
	ببنيتها, 26				بنية	213117	2	ب
	ببنيتك, 22				بنية	213117	2	ب
	واللبنية, 20				بنية	213117	2	ب
	وببنية, 19				بنية	213117	2	ب
	ببنياتها, 19				بنية	213117	2	ب
	وبالبنية, 19				بنية	213117	2	ب
	البنيتان, 18				بنية	213117	2	ب
	وبنيتها, 16				بنية	213117	2	ب
	ولبنية, 16				بنية	213117	2	ب
	فبنيتها, 16				بنية	213117	2	ب
	فبنيتها, 14				بنية	213117	2	ب
	لبنياته, 14				بنية	213117	2	ب

الرمادي	الرمادي	الرمادي	الرمادي	الرمادي	الرمادي	
6616	11629	الرمادية، 1340 ورمادية، 64 والرمادية، 57 الرماديين، 31 الرماديتان، 17 الرماديين، 15 الرماديات، 9 الرماديون، 7 بالرمادية، 5 الرمادية، 5 برمادية، 4 والرماديات، 4 الرماديان، 3 فرمادية، 3 ورمادية، 2 ورماديته، 1 لرمادية، 1 الرمادي، 1 والرماديين، 1 والرماديون، 1 للرماديات، 1 ورماديين، 1 ورماديتها، 1 برماديات، 1 كالرماديات، 1 الرماديكما، 1 برماديتها، 1	رمادية، 1113 رماديين، 13 رماديين، 12 رماديه، 10 رماديتان، 7 رماديان، 6 رماديات، 5 رمادية، 4 رماديون، 3 رماديتها، 2 رمادياتها، 1	الرمادي، 6616 والرمادي، 1048 بالرمادي، 113 للرمادي، 67 ورمادي، 53 كالرمادي، 14 فالرمادي، 12 برمادي، 2 لرمادي، 2 الرمادي، 2 وللرمادي، 1	3 4	945 0
الزيتوني الزيتوني	1269	الزيتونية، 170 الزيتونيين، 39 والزيتونية، 5 بالزيتونية، 3 الزيتونه، 3 بالزيتونيين، 1 والزيتونيين، 1 الزيتونيين، 1 لزيتونيين، 1	زيتونية، 22 زيتونيين، 4 زيتونيان، 1 زيتونيون، 1	الزيتوني، 490 والزيتوني، 16 بالزيتوني، 9 وزيتوني، 8 للزيتوني، 3 لزيتوني، 1	3 4	490 0
الزيتوني						

معاجم الشبوع والمدونات اللغوية العربية
دراسة على ضوء لسانيات المدونات

الوردي	الوردي	الوردي	الوردي	الوردي
5175	8873	2011	1496	5175
		الوردي	وردي	الوردي
		229	207	والوردي
		كوردي	ورديتين	كوردي
		69	30	ووردي
		وردي	ورديته	32
		64	13	بالوردي
		42	6	29
		الورديان	ورديتهم	بوردي
		37	5	16
		الورديتين	ورديان	كالوردي
		35	3	أوردي
		للوردي	ورديتهم	14
		27	3	لوردي
		لوردي	ورديهم	10
		15	2	فوردي
		11	2	8
		كورديان	ورديته	فالوردي
		11	2	6
		أكورديون	ورديته	وبالوردي
		الورديتان	ورديته	3
		10	2	ولوردي
		كورديه	ورديته	1
		9	1	وللوردي
		الورديين	ورديته	1
		8	1	وكوردي
		بالورديات	ورديته	
		7	1	
		بورديات	ورديته	
		7	1	
		والورديات	ورديته	
		6		
		بورديه		
		6		
		والورديان		
		5		
		بورديه		
		4		
		أورديه		
		3		
		وكورديه		
		2		
		للورديات		
		2		
		فورديه		
		2		
		أوكورديون		
		2		
		بورديك		
		2		
		بالورديان		
		2		
		لورديه		
		2		
		أورديه		
		2		
		لورديات		
		2		
		وورديات		
		2		
		ووردياته		
		1		
		كوردي		
		1		
		بوردياته		
		1		
		بوردياتهم		
		1		
		كالورديان		
		1		
		ووردياتها		
		1		
		أورديين		
		1		
		كورديات		
		1		
		كورديته		
		1		
		بورديون		
		1		
		بورديان		
		1		
		والورديتان		
		1		
		لورديان		
		1		
		بالورديون		
		1		
		أوورديات		
		1		
		فورديه		
		1		
		والورديون		
		1		
		وورديته		
		1		
		فالورديه		
		1		

مجموع التكرارات لجميع المفردات

615757

8- المراجعة السياقية لتكرارات كل مصطلح باستخدام الكشاف السياقي: 1/8 - قيم الإحصاء بعد المراجعة السياقية:

تبين في مرحلة البحث الآلي تباين التكرارات بين البحث عن تكرار جذع المصطلح بالشكل، والبحث بدون الشكل، وفي هذه المرحلة قمنا بالبحث باستخدام الكشاف السياقي عن كل جذع مع كل الصور المنبثقة عنه كذلك. وقد حددنا البحث بعدد ثماني كلمات قبل كلمة البحث، و ٨ كلمات بعدها. وفي هذه المرحلة قمنا بمراجعة سياقات ما يزيد عن ستمئة وخمسة عشر ألف تكرار، تحديداً راجعنا (٦١٥٧٥٧) ستمئة وخمسة عشر ألفاً، وسبعمئة وسبعة وخمسين سياقاً لكل تكرار.

وقد أسفرت المراجعة عن استبعاد تكرارات كثيرة، فتغيرت قيم الإحصاء الأولية لمصطلحات العلوم الأساسية في المدونة العربية بشكل واضح، وتباينت القيم بين الإحصاء الآلي قبل المراجعة السياقية، والإحصاء الآلي بعد المراجعة السياقية. والجدول التالي يوضح القيم الإحصائية.

مصطلح اللون الأساسي	تكرار جذع المصطلح مع الشكل	تكرار جذع المصطلح بدون شكل	مجموع تكرارات المصطلح قبل المراجعة السياقية	مجموع تكرارات المصطلح بعد المراجعة السياقية	الصورة الأكثر استعمالاً	تكرار الصورة الأكثر استعمالاً قبل المراجعة السياقية	تكرار الصورة الأكثر استعمالاً بعد المراجعة السياقية
أحمر	١١٨	١٣٨٨٧	٧٠٠٦٧	٦٧٩٦٤	الأحمر	٤٩٧٠٨	٤٩٢١٠
أسود	٢٠٩	١٨٩٢٠	٨٠١٣١	٦٤١٠٤	الأسود	٥٠٥٥٦	٤٠٤٤٥
أبيض	٥٦	١٧٤٥١	٦٣١٨١	٦٠٠٢١	الأبيض	٤٠١٨٨	٣٩٧٨٦
أخضر	٥١	٧٥٩٨	٣٧٥٦٨	٣٧١٩٢	الأخضر	٢٦٥٨١	٢٦٣١٥
أزرق	٤	٣١٢٧	٢١٤٩٠	٢١٢٧٥	الأزرق	١٥٠٢٧	١٤٨٧٦
أصفر	٢٣	٣٨٨٤	١٣٨٩٤	١٣٧٥٥	الأصفر	٧٨٩٦	٧٨١٧
رمادي	٠	٩٤٥	١١٦٢٩	١١٥١٢	الرمادي	٦٦٦٦	٦٥٤٩
وردي	٥	١٢٩٤	٨٨٧٣	٥٧٦٧	الوردي	٥١٧٥	٣٣٦٣
أسمر	٥	٢٠٠١	٧٤٧٣	٤٨٥٧	الأسمر	٢٧٠٩	١٧٦٠
برتقالي	٠	٣٠٥	٤٣٣٥	٤٢٩١	البرتقالي	٢٣٩٧	٢٣٧٣
بنفسجي	٠	٣٨٨	٤٢٢٨	٤٢٢٨	البنفسجي	٢٣٠١	٢٣٠١
بي	٢	٢١٢١١٧	٢٩٠٠٥٤	١١٦٠٢	البيضاء	٣٨٣٣٨	١٥٣٣
أرجواني	٠	١١٢	٨٨٧	٨٦٩	الأرجواني	٤٠١	٣٩٢
زيتوني	٠	٤٩٠	١٢٦٩	١٩٠	الزيتوني	٤٩٠	٩٨
أدكن	٢	١٠٦	١٧٧	١٠٢	أدكن	١٠٦	٦٢
أزود	٤	٣٧٧	٥٠١	٢	أزود	٣٧٧	٢
المجموع	٤٧٩	٢٨٣٩٠٢	٦١٥٧٥٧	٣٠٧٧٣١		٢٤٨٨٦٦	١٩٦٨٨٢
أكبر قيمة	٢٠٩	٢١٢١١٧	٢٩٠٠٥٤	٦٧٩٦٤		٥٠٥٥٦	٤٩٢١٠
أصغر قيمة	٠	١٠٦	١٧٧	٢		١٠٦	٢

جدول ١: قيم إحصاء شيوخ مصطلحات الألوان الأساسية.

2/8- النتائج الإحصائية بعد المراجعة السياقية:

أولاً- نتائج الإحصاء الوصفي:

يمكن من الجدول أن نحصل من القيم الإحصائية على نتائج الإحصاء الوصفي التالية:

- ١- مجموع تكرارات جذع المصطلح مع الشكل بلغت تسعة وسبعين وأربعمائة ٤٧٩ تكرار، وكانت أعلى قيمة تكرار للون الأصفر وهي: مئتان وتسعة ٢٠٩ تكرارات. وأقل قيم التكرار كانت للرمادي، والبرتقالي، والأرجواني، والزيتوني بصفر تكرارات.
- ٢- مجموع تكرارات جذع المصطلح بدون شكل بلغت مئتين وثلاث وثمانين ألفاً واثنين وتسعمائة مصطلح. وكانت أعلى قيمة تكرارات للون البني وهي مئتان وثلاثة عشر ألفاً ومئة وسبعة عشر تكراراً. وأقل قيمة للتكرار كانت للون الأدكن، وهي مئة وستة ١٠٦ تكرارات.
- ٣- مجموع تكرارات المصطلح قبل المراجعة السياقية بلغت ستمئة وخمسة عشر ألفاً وسبعمئة وسبعة وخمسين تكراراً، وكانت أعلى قيمة لتكرار اللون البني، وهي مئتان وتسعون ألفاً وأربعة وخمسون تكراراً. وأقلها تكراراً للون الأدكن، وهي مئة وسبعة وسبعون ١٧٧ تكراراً.
- ٤- مجموع تكرارات المصطلح بعد المراجعة السياقية بلغت ثلاثمئة وسبعة آلاف وسبعمئة وواحد وثلاثين تكراراً. وكانت أعلى التكرارات للون الأحمر بقيمة تكرار هي سبع وستون ألفاً وتسعمئة وأربعة وستون تكراراً، وكانت أقل التكرارات للون الأرمد بقيمة تكرارين فقط.
- ٥- مجموع تكرارات الصورة الأكثر شيوعاً للمصطلح قبل المراجعة السياقية بلغت مئتين وثمانين وأربعين ألفاً وثمانمئة وستة وستين تكراراً. وكان أعلى تكرار للون الأسود، وهو خمسون ألفاً وخمسمئة وستة وخمسون تكراراً، وأقل قيمة تكرار كانت للون الأدكن، وهو مئة وستة تكرارات.
- ٦- مجموع تكرارات الصورة الأكثر شيوعاً للمصطلح بعد المراجعة السياقية بلغت مئة وستة وتسعين ألفاً وثمانمئة واثنين وثمانين تكراراً. وكان أعلى تكرار للون الأحمر، وهو تسعة وأربعون ألفاً ومئتين وعشرة تكرارات، وأقل قيمة تكرار كانت للون الأرمد، وهو تكراران فقط.
- ٧- المتوسط الحسابي لتكرار جذع المصطلح المشكول وغير المشكول:

المتوسط الحسابي	
٢٩.٩٤	تكرار جذع المصطلح مع الشكل
١٧٧٤٣.٨٧	تكرار جذع المصطلح بدون شكل

٨- المتوسط الحسابي لتكرار المصطلح قبل المراجعة السياقية وبعد المراجعة السياقية:

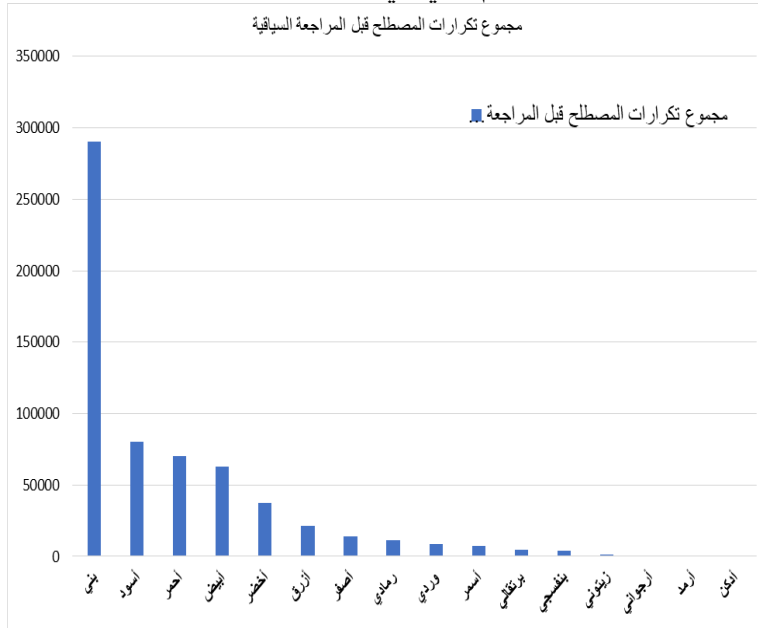
المتوسط الحسابي	
38484.81	تكرارات المصطلح قبل المراجعة السياقية
19233.18	تكرارات المصطلح بعد المراجعة السياقية

٩- المتوسط الحسابي لتكرار الصورة الشائعة للمصطلح قبل المراجعة السياقية وبعد المراجعة السياقية:

المتوسط الحسابي	
15554.12	تكرار الصورة الأكثر استعمالاً قبل المراجعة السياقية
12305.12	تكرار الصورة الأكثر استعمالاً بعد المراجعة السياقية

١٠- ترتيب مصطلحات الألوان الأساسية وفق الشبوع قبل المراجعة السياقية:

جاء اللون البني متصدرًا ترتيب التكرارات، يليه اللون الأسود، فالأحمر، فالأبيض، ثم يأتي في ذيل القائمة اللون الأدكن.

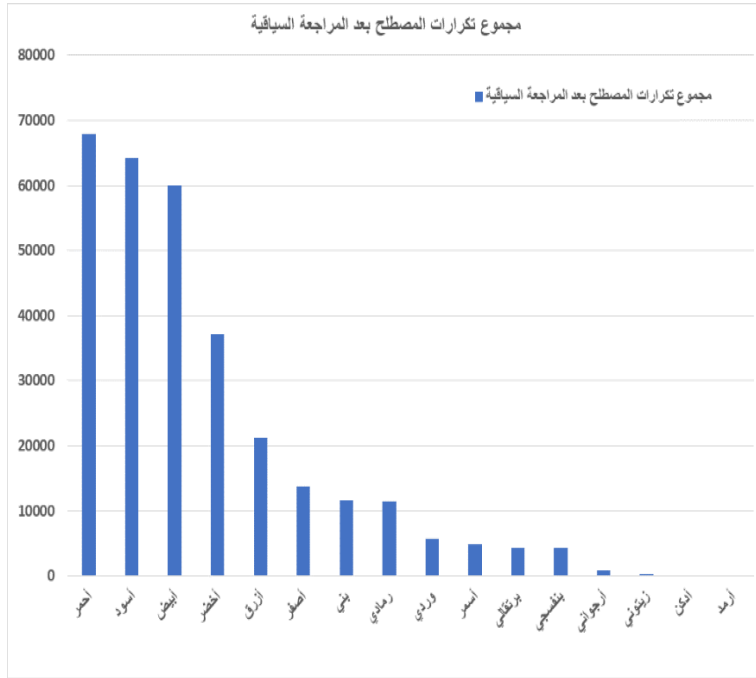


شكل ٨: مجموع تكرارات مصطلحات الألوان الأساسية قبل المراجعة السياقية

١١- ترتيب مصطلحات الألوان الأساسية وفق الشبوع بعد المراجعة السياقية:

جاء اللون الأحمر في صدارة شبوع الاستعمال، يليه اللون الأسود، ثم الأبيض. وكانت أقل الألوان شبوعاً في الاستعمال هي الأدكن ثم الأرمو.

معاجم الشبوع والمدونات اللغوية العربية دراسة على ضوء لسانيات المدونات

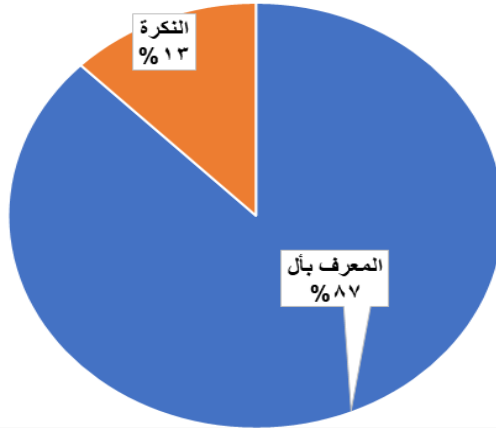


شكل ٩: مجموع تكرارات مصطلحات الألوان الأساسية بعد المراجعة السياقية

١٢- الصور الأكثر شيوعاً في الاستعمال لمصطلحات الألوان الأساسية:

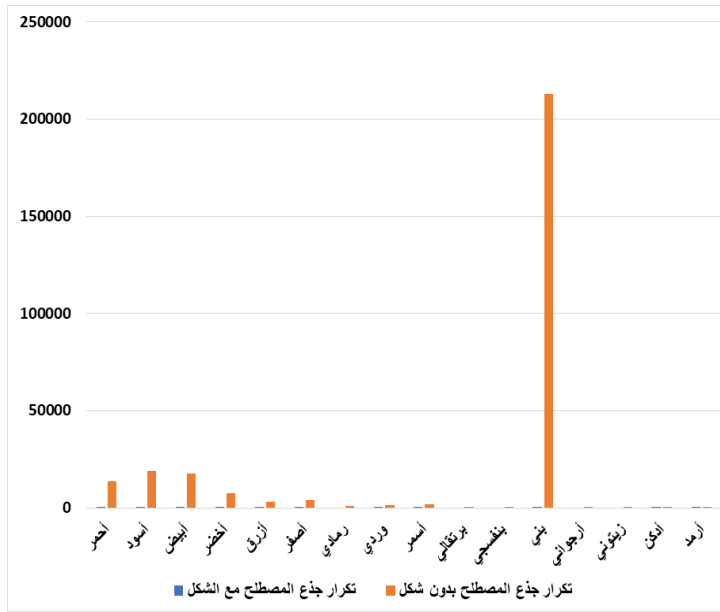
كان مصطلح "الأحمر" بأل التعريف هو صورة اللون الأكثر شيوعاً من بين الصور الشائعة لمصطلحات الألوان يليه مصطلح "الأسود" بأل التعريف، وجاء في ذيل القائمة اللون "أرمد" بدون أل التعريف، يسبقه في ذيل القائمة اللون "أدكن" بدون أل التعريف. واللافت أن جميع صور مصطلحات الألوان الشائعة كانت معرفة بأل ما عدا "أرمد" و"أدكن".

الصورة الأكثر شيوعًا لمصطلحات الألوان الأساسية



شكل ١٠: الصورة الأكثر شيوعًا لمصطلحات الألوان الأساسية

١٣- تكرار جذع المصطلح المشكول كانت في كل مرة أقل من جذع المصطلح غير المشكول. ويمكن توضيح ذلك بالشكل التالي:

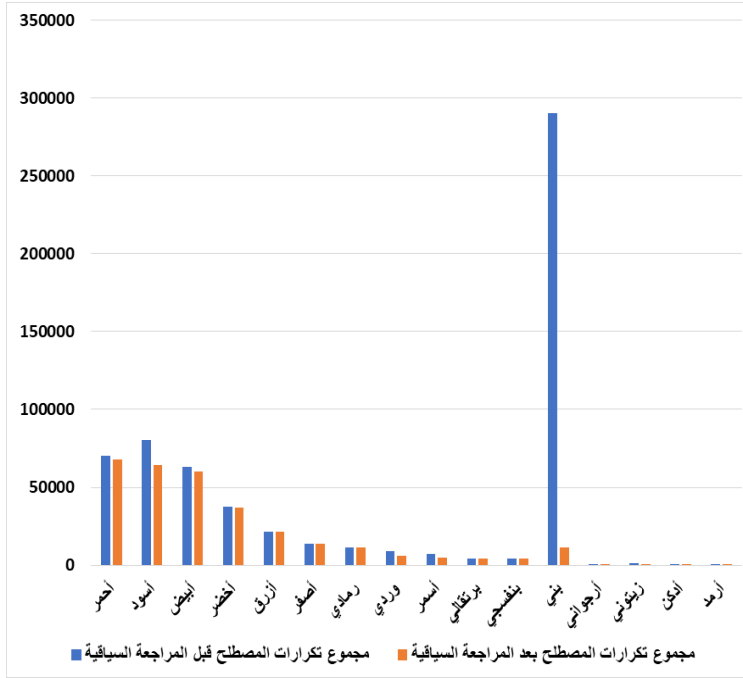


شكل ١١: مقارنة تكرارات جذع المصطلح مع الشكل وبدون شكل

١٤- مجموع تكرارات المصطلح الآلي قبل المراجعة السياقية أكبر في كل مرة من مجموع التكرارات بعد المراجعة السياقية، ما عدا في مصطلح اللون البنفسجي

معاجم الشبوع والمدونات اللغوية العربية دراسة على ضوء لسانيات المدونات

فقط، حيث تساوى شبوع المصطلح في المرتين، ويمكن توضيح ذلك بالشكل التالي:



شكل ١٢: مقارنة تكرارات المصطلح قبل المراجعة السياقية وبعدها.

ومن هذه النتيجة الأخيرة يتبين أن مصطلح بنفسي هو مصطلح اللون الأساسي الوحيد الذي لا يستخدم في دلالة أخرى غير دلالاته على اللون، ولا يلتبس في استعماله مع مفردات أخرى. وهذا الكلام يصدق على الجذع بنفسي، وعلى كل ما انبثق عنه من صور.

كذلك يمكن أن نستنتج أن الاعتماد على البحث الآلي لا يتسم بالدقة الكافية، وقد يؤدي لإحصاءات مضللة، ولا يمكن الاستغناء في الوضع الحالي للمدونات العربية- عن المراجعة السياقية للنصوص. فعلى سبيل المثال كلمة أرمو أحصاها الباحث الآلي ما يزيد عن خمسمئة مرة، لكن الكشف السياقي اليدوي أثبت من هذا العدد مرتين فقط استخدم فيهم المصطلح في دلالاته على اللون وهو ما يعادل أربعة في الألف ٠.٤% من التكرار السابق. ومصطلح بني الذي زادت تكراراته في الإحصاء الآلي عن مئتين وتسعين ألف تكرر، وبعد الكشف السياقي تبين أن هذا المصطلح دل على اللون فيما يقل عن أربعة بالمائة ٤% من هذا الرقم فقط، حيث ورد في اثنين وستمئة وأحد عشر ألف تكرر فقط.

ثانيًا- النتائج الإحصاء التحليلي:

سنحاول في هذا الجزء معرفة ما يمكن أن تعطيه قيم الإحصاء من نتائج تحليلية، سواء على مستوى الارتباط الدال إحصائيًا، أو الفروق الدالة إحصائيًا، وهل يمكن استنتاج أي علاقة إحصائية بين قيم تكرارات مصطلحات الألوان الأساسية. واستخدمنا في إعداد النتائج التحليلية برنامج الإحصاء (IBM SPSS Statistics Base 20.0).

١- الارتباط بين إحصاء جذع المصطلح مع الشكل وبدون الشكل.

أسفر تحليل الارتباط باستخدام معامل بيرسون عن النتائج التالية:

تكرار جذع المصطلح بدون شكل	تكرار جذع المصطلح مع الشكل		
-0.022- 0.936 16	1 16	معامل الارتباط (بيرسون) دالة الارتباط عدد المصطلحات	تكرارات جذع المصطلح مع الشكل
1 16	-0.022- 0.936 16	معامل الارتباط (بيرسون) دالة الارتباط عدد المصطلحات	تكرارات جذع المصطلح بدون شكل

جدول 2: جدول الارتباط بين إحصاء جذع المصطلح مع الشكل، وبدون الشكل.

وهذا يعني أنه لا يوجد ارتباط دال إحصائيًا بين إحصاءات جذع المصطلح مع الشكل، وإحصاءاتها بدون الشكل عند مستوى دلالة 0.05.

٢- الارتباط بين مجموع تكرارات المصطلح قبل المراجعة السياقية وبعد المراجعة السياقية:

أسفر تحليل الارتباط باستخدام معامل بيرسون عن النتائج التالية:

تكرار جذع المصطلح بعد المراجعة السياقية	تكرار جذع المصطلح قبل المراجعة السياقية		
0.287 0.282 16	1 16	معامل الارتباط (بيرسون) دالة الارتباط عدد المصطلحات	تكرارات المصطلح قبل المراجعة السياقية
1 16	0.287 0.282 16	معامل الارتباط (بيرسون) دالة الارتباط عدد المصطلحات	تكرارات المصطلح بعد المراجعة السياقية

جدول 3: جدول الارتباط بين مجموع تكرارات المصطلح قبل المراجعة السياقية، وبعد المراجعة السياقية.

وهذا يعني أنه لا يوجد ارتباط دال إحصائيًا بين إحصاءات المصطلح قبل المراجعة السياقية، وإحصاءات المصطلح بعد المراجعة السياقية عند مستوى دلالة 0.05.

٣- الفروق بين إحصاء جذع المصطلح مع الشكل وبدون الشكل.

أسفر تحليل الفروق باستخدام اختبار العينات المقترنة عن النتائج التالية:

معاجم الشبوع والمدونات اللغوية العربية
دراسة على ضوء لسانيات المدونات

اختبار فروق العينة المقترنة

	فروق العينة المقترنة				إحصائية ت	درجة حرية الإحصاء df	الدالة (الطرفين) Sig. (2-tailed)
	الاحرف المعياري	متوسط الخطأ المعياري	%٩٥				
			الأعلى	الأدنى			
تكرار جذع المصطلح مع الشكل زوج ١ تكرار جذع المصطلح بدون شكل زوج ١	١٧٧١٣,٩٢٨٠	٥٢٤٩٠,٣٠٠	١٣١٢٢,٥٧٥	٤,٠٤٤٤٤	١٠٢٥٦,١٦٩	١٥	٠,١٩٧

جدول ٤ : جدول الفروق بين إحصاء جذع المصطلح مع الشكل، وبدون الشكل

وتبين النتائج أنه لا توجد دلالة إحصائية للفروق بين جذع المصطلح المشكول، وجذع المصطلح غير المشكول عند مستوى دلالة 0.05.

٤- الفروق بين مجموع تكرارات المصطلح قبل المراجعة السياقية، وبعد المراجعة السياقية:

أسفر تحليل الفروق باستخدام اختبار العينات المقترنة عن النتائج التالية:

اختبار فروق العينة المقترنة

	فروق العينة المقترنة				إحصائية ت	درجة حرية الإحصاء df	الدالة (الطرفين) Sig. (2-tailed)
	الاحرف المعياري	متوسط الخطأ المعياري	%٩٥				
			الأعلى	الأدنى			
مجموع تكرارات المصطلح قبل المراجعة السياقية زوج ١ مجموع تكرارات المصطلح بعد المراجعة السياقية زوج ١	١٩٢٥١,٦٦٥	٦٩٢٣١,٤٥٨	١٧٣٠٧,٨٦٤	-١٧٢٣٩,٦١٥	٥٦١٤٦,٤٦٥	١٥	٠,٢٨٤

جدول ٥: جدول فروق مجموع تكرارات المصطلح قبل المراجعة السياقية، وبعد المراجعة السياقية.

وتبين نتائج الإحصاء التحليلي أنه لا توجد دلالة للفروق الإحصائية بين تكرارات المصطلح قبل المراجعة السياقية، وبعد المراجعة السياقية عند مستوى دلالة 0.05.

خلاصة: توضح القيم الإحصاء الوصفي أن هناك تباينات كبيرة بين تكرارات المصطلح المشكول، والمصطلح غير المشكول، وبين تكرارات المصطلح قبل المراجعة السياقية، وبعد المراجعة السياقية.

يبين التحليل الإحصائي للارتباط أنه لا يوجد ارتباط يمكن أن يعطينا دلالة إحصائية مفيدة بين تكرارات جذع المصطلح المشكول، والجذع غير المشكول، وبالمثل لا يوجد رابط يمكن أن يعطينا دلالة إحصائية مفيدة بين تكرارات المصطلحات قبل المراجعة السياقية، وبعد المراجعة السياقية. فأرقام الارتباط لا تقدم أية نتائج إحصائية يبني عليها التحليلات الماثلة لاحقاً.

كذلك يبين التحليل الإحصائي للفروق أنه لا توجد دلالة إحصائية للفروق بين إحصاء جذع المصطلح المشكول وإحصاء الجذع غير المشكول، وكذلك

لا يوجد دلالة إحصائية للفروق بين إحصاءات شيوع المصطلح قبل المراجعة السياقية وبعد المراجعة السياقية. أي أنه لا توجد علاقة فروق إحصائية يبني عليها التحليلات المماثلة لاحقاً.

٣/٨- إشكاليات القياس الآلي لشيوع مصطلحات الألوان الأساسية في المدونة العربية.

١/٣/٨- عدم ضبط بنية الكلمات بالشكل:

في التدوين العربي يرمز للشدة والسكون والصوائت القصيرة بخمسة رموز توضع فوق أو تحت الصامت. هذه الرموز هي: [الفتحة (-)، والضمة (ـُ)، والكسرة (-ِ)، والسكون (-َ)، والشدة (-)]. وهذه الرموز تستعمل لتوضيح طريقة نُطق الكلمة. وهذه الرموز مهمة جداً في التدوين الكتابي للتفريق بين الكلمات متشابهة الصوامت، مثل (كُتِبَ) و (كُتِبَ) و (كُتِبَ). كذلك تستعمل رموز الصوائت القصيرة بالإضافة لرموز التنوين [تنوين النصب (-ِ)، تنوين الرفع (-َ)، تنوين الجر (-ِ)] علامة على الإعراب، ومحلها الصامت الأخير من المفردة.

لكن في التدوين الكتابي المعاصر أصبح استخدام هذه الرموز في شكل بنية المفردة مهماً إلى حد كبير؛ نظراً لاعتماد المتحدث العربي الأصلي على سياق الكلمة لتوضيح المعنى، ومن ثم استنتاج النطق الصحيح للمفردات دون حاجة ملحّة في أغلب الأحيان- لهذه الرموز. وهذه الأمر مثّل لصانعي المدونات إشكالية مهمة، لأن أغلب النصوص التي تجمعها المدونات تكون نصوصاً خالية من شكل البنية.

وقد تبين من النتائج الأولية لقياس شيوع الآلي لمصطلحات الألوان الأساسية أن إشكالية عدم شكل بنية المفردات في المدونة كانت بارزة بوضوح، وذلك كما يلي:

١- جميع المصطلحات -بلا استثناء- وجد بها فروق إحصائية بين تكرار المصطلح بالشكل الكلي، وتكرار المصطلح بغير الشكل الكلي.

٢- تبين أن قياس التكرار آلياً في غياب الشكل الكلي جاء بنتائج غير دقيقة في كل المصطلحات. حيث نجم عن مراجعة السياق لكل تكرار من التكرارات كشف أخطاء كثيرة بسبب إحصاء مفردات ليس لها علاقة بالمصطلح، لكنها تتشابه معه في الصوامت دون الصوائت. وهذه مشكلة حلها الأول هو شكل بنية الكلمة. ولننظر في الأمثلة التالية:

ورد في إحصاء مصطلح "أسود" التكرارات التالية:

- "كذلك الحيوانات أسود ضارية مفترسة شرسة،
ونسور ذات بأس شديد".
- "العرب لا تغفو، ولا يغفون أمثالك من أسود
الفكر".
- وبما أن ال أسود تأكل وجبة واحدة يومياً فقد قرر
الرجل تقليدها.
- وهنا نجد في الأمثلة السابقة أن المقصود هو "أسود" جمع أسد،
وليس اللون الأسود. وتم إحصاؤها ألياً في تكرارات مصطلح أسود
لتشابه التدوين. وشكل بنية الكلمات حل مناسب لعلاج هذا الخطأ
الإحصائي.
- كذلك ورد في إحصاء مصطلح "بني" التكرارات التالية:
- بسبب استهدافه لشيخ كبير في احدى غاراته على
بني نبهان، فوافاه بطل منهم بطعنة... .
- من المشكلات التي تواجه ساكنيه، أي نحن بني
البشر. إذ يحتوي الموقع على عدادات تمثل... .
- وقد نشأ الفن الاسلامي في دمشق مع عصر بني
أمية، وكان الطراز الأموي هو الذي يحتذى.
- وقد تركت الشائعات أثرها أيضا على سهم بني
سويف بعد أن تأخر إلى هذه اللحظة توثيق.
- ... المتواجدة حالياً في ميناء بني غازي في ليبيا،
والتي يريد منظموها التوجه.
- في عهد دولة بني العباس.
- في هذه الأمثلة المذكورة يوضح السياق أن المقصود من المفردة
هو "بني" جمع ابن، التي تستخدم في معناها المباشر مثل (بني
البشر) أو استخدمت في شراكة مع اسم علم من الأعلام (بني سويف،
بني غازي، بني نبهان، بني أمية، بني العباس). وعلى الرغم من ذلك
تم إحصاء عشرات من هذه المفردة (بني) ألياً في تكرارات
المصطلح "بني" لتشابه التدوين. وعلاج هذا الخطأ في الإحصاء
الألي يتطلب التشكل الكلي لبنية الكلمة.
- وفي إحصاء المصطلح "بني" أيضاً نجد الأمثلة التالية:
- نسميه بالاستعراض السياسي على حساب هيكل
سياسي متماسك بني على تراكم خبرة وحكمة قيادة ذات
نظرة ثاقبة.

- ... ضخم من المصانع ومختبرات البحث والمخابئ والمرافق الأخرى بُني على أرض جدباء تقع على بعد ٥٠ ميلاً.
- الصيام ٩٠١ - (حديث ابن عمر: بني الإسلام على خمس).
- كما بُني للمسجد أروقة.
- في الأمثلة السابقة يكشف السياق أن المقصود هو الفعل "بُنِيَ" وهو المبني للمفعول من الفعل "بَنَى"، ولا يمكن أن يكون المقصود هو المصطلح "بُنِيَ" الدال على اللون. وتم إحصاء تكرارات هذه المفردة مع تكرارات مصطلح اسم اللون لتشابه التدوين. وعلاج ذلك يكمن في ضبط البنية بالشكل. كذلك في تكرارات مصطلح "بني" وردت الأمثلة التالية:
- فقال: يا بني، إني لم أرد بإخراجك له طرده، ...
- قال بعض أئمة السلف لأحد المؤذنين: يا بني: إنك ترقى على المسجد، فإذا رقيته ...
- في هذين المثالين يكشف السياق أن المقصود هو "بُنِيَ" تصغير كلمة "ابن"، وليس المقصود هو المصطلح "بُنِيَ" الدال على اللون، وعلى الرغم من ذلك تم إحصاء "بُنِيَ" ألياً في تكرارات مصطلح اسم اللون لتشابه التدوين. وحل هذه الإشكالية يتطلب الشكل الكلي لبنية مفردات المدونة. وفي تكرارات مصطلح "بُنِيَ" أيضاً وردت الأمثلة التالية:
- عبر مقابلة طويلة أجراها مع المؤرخ الإسرائيلي "بني موريس" ...
- ابن الحسن أبو السعود المعروف بابن الـبني بقراءتي عليه ببغداد، ...
- في كلا المثالين نجد أن المقصود اسم علم على إنسان هو "بُنِيَ"، وليس مصطلح اسم اللون. ولكن الإحصاء الآلي جمع بالخطأ تكرارات هذه المفردة مع تكرارات جذع مصطلح اسم اللون لتشابه التدوين. ولمنع تكرار هذا الخطأ في الإحصاء الآلي لابد من الشكل الكلي لبنية مفردات المدونة.
- وفي تكرارات مصطلح "بُنِيَ" وردت الأمثلة التالية:
- ... والإدارات المعنية بتأسيس البنية التحتية لهذا المشروع ...
- واستعراض أهمية البنية التحتية للنمو الاقتصادي المستقبلي.
- ابن الـ ٢٢ ذو البنية الجسدية المتواضعة ...

- ... وكيف خضعت البنية الاجتماعية لهذه النظريات.
- ... لا بد من إتقانها لإيجاد البنية المعلوماتية اللازمة.
- ... فولهام في منتصف الملعب مستغلاً البنية الجسمانية التي يتمتع بها لاعبوها.
- ... لمختلف قطاعات الجامعة، استكمال البنية الأساسية لمراكز البحوث العلمية، ...
- في الأمثلة السابقة تكشف السياقات أن المقصود هو لفظ "البنيّة" وهو مشتق من الفعل "بنى"، وليس له علاقة بالمصطلح "البنيّة" الدال على اللون. وعلى الرغم من ذلك أدخله الإحصاء الآلي في تكرارات جذع مصطلح اسم اللون "بني". والسبب في ذلك هو عدم الشكل الكلي لبنية مفردات المدونة. وفي تكرارات مصطلح "وَرْدِيّ"، ورد المثال التالي:
- وردِيّ الوحيد علي شعلان بخصوص اتهامه لي بعدم الاهتمام بسرقة لوحة ...
- وفي هذا المثال نرى أن السياق يكشف عن المقصود هو "وَرْدِيّ"، أي الفعل "رَدَّ" مسبق بواو العطف، ومسند إلى ياء المتكلم (ي)، وليس المقصود قطعاً مصطلح اسم اللون "وَرْدِيّ". ومع ذلك جمع الإحصاء الآلي تكرارات هذه المفردة في تكرار مصطلح اسم اللون "وردي". وهذا الخطأ يمكن اتقاؤه بالشكل الكلي لمفردات المدونة. وفي تكرارات مصطلح "أرْمَد"، ورد المثال التالي:
- ... في تقرير نشرته مجلة "أرمد فورسز" الأمريكية العسكرية المتخصصة في يوليو ...
- في هذا المثال يكشف السياق أن "أرْمَد" هي تعريب لاسم مجلة أمريكية، ومعناه "عسكري"، ولا علاقة لهذه الكلمة بمصطلح اللون الأساسي "أرْمَد". ونظراً لنتشابه المفردتين تدويناً أدخل الإحصاء الآلي تكرارات اسم العلم على مكان، مع تكرارات مصطلح اسم اللون. وكان يمكن بالشكل الكلي أن يصحح هذا الخطأ.
- إذن بدون شكل كلي لمفردات المدونة، فإن دقة الإحصاء الآلي للمفردات يتطلب المراجعة البشرية اليدوية. فالآلة -حتى الآن- قدرتها لا تقارن بقدرة العقل البشري على فهم الدلالة، واستقراء السياقات، واختيار الكلمة المناسبة للنص دون ضرورة ملحة للشكل الكلي. لهذا فإن الباحث الآلي في النصوص العربية إذا غاب الشكل الكلي لبنية جميع المفردات لن يتمتع بكفاءة كبيرة عند البحث الآلي في التمييز بين المفردات المتشابهة تدويناً كتابياً، والمختلفة نطقاً ودلالة. وللتغلب على هذه الإشكالية نوصي بأن تقوم إدارة المدونة العربية

بالتخطيط لشكل بنية جميع مفردات المدونة شكلاً كلياً، ويعتمدون في ذلك على مرحلتين:
المرحلة الأولى: هي الشكل الكلي للبنية آلياً. باستخدام برامج مثل برنامج شَكَلِي من إنتاج شركة صخر (<https://tashkeel.alsharekh.org>).
المرحلة الثانية: هي مراجعة الشكل الكلي الآلي يدوياً بالاستعانة بمجموعة كبيرة من المدونين ذوي الخبرة.
مع تقبل حقيقة أن نسبة دقة أي من هاتين الطريقتين قد لا يصل إلى نسبة ١٠٠%.

٢/٣/٧ - أخطاء التدوين.

من الإشكاليات التي أثرت في دقة الإحصاء الآلي للشبوع وجود عدد من أخطاء التدوين في المفردات، ومن أمثلة ذلك:
في تكرارات مصطلح اسم اللون "بُنِّي"، وردت الأمثلة التالية:
○ مثلي ومثل الأنبياء من قبلي كمثل رجل بُني بيتاً فأجمله وأحسنه إلا موضع لبنة ...
هنا السياق يكشف أن المقصود هو الفعل (بني)، لكن حدث خطأ في تدوين الألف اللينة (ى) حيث دونت بالياء (ي)، فكتب الفعل خطأ بهذا الشكل (بني). وترتب على ذلك أن تدخل تكرارات الفعل (بني) في تكرارات مصطلح اسم اللون الأساسي "بُنِّي". وعلاج ذلك هو دقة التدوين الكتابي، والتنبه للحروف المتشابهة تدويناً، مثل (ى ، ي)، و (أ ، ا ، أ)، و(ه ، ة) وهكذا.
كذلك في تكرارات مصطلح اسم اللون الأساسي "بُنِّي" في صورته المعرّفة "البُنِّي" وردت الأمثلة التالية:
○ هو الذي خاصم بني عقيل إلى البني صلى الله عليه وسلم في العقيق فقضى ...
○ فقال: لما قال البني صلى الله عليه وسلم: "من كنت ..."
○ ... الخندق فما بعدها، وكان عامل البني -صلى الله عليه وسلم- على نجران، ...
○ ... طول العمر في طاعة الله، وقال البني صلى الله عليه وسلم: "إذا بلغ المرء المسلم ..."
○ ... أهل المدينة: إنه قدم على البني -صلى الله عليه وسلم- من مكة فقالت عائشة: ...
○ ... بن مالك عن أبيه: عن البني -صلى الله عليه وسلم- أرواح الشهداء في ...

○ ... البخاري، عن عائشة، عن النبي -صلى الله عليه وسلم- أنه قال: ...

وفي جميع الأمثلة السابقة يكشف السياق بوضوح أن المقصود هو كلمة "النَّبِيَّ"، وليس مصطلح "النَّبِيَّ". وهذا الخطأ في التدوين متكرر بشكل كبير جداً مما يرجح -لدي- أن ثمة استبدالاً كلياً حدث -بالخطأ وعن غير قصد- لهذه الكلمة في مصدر من مصادر المدونة العربية. وعلى الرغم من وضوح هذا الخطأ لعين الفاحص البشري، إلا أن الباحث الآلي لم يتنبه لخطأ التدوين، وجمع تكرارات هذه الكلمة الخطأ تدوينياً مع تكرارات مصطلح اسم اللون الأساسي "النَّبِيَّ".

٨/٣-٣. اتحاد اللفظ، واختلاف المعنى.

يوجد بالعربية بعض الكلمات التي تتحد في اللفظ والبنية الصرفية، لكنها تختلف في المعنى. ومع مثل هذه الكلمات يكون دور السياق حاسماً في تحديد دلالة المفردة. كذلك نجد أن الإحصاء الآلي لا يكون دقيقاً في فصل تكرارات المفردة المقصودة عن بقية التكرارات غير المقصودة. ولننظر في الأمثلة التالية:

○ ويقال لا أفعله ما أسمر بنا سмир. والسمير الدهر.
هنا المفردة الواردة هي "أسمر". ويكشف السياق أنها فعل ماض مضارعة يسمر، أي يتحدث ليلاً، وفلانٌ سَمِيرٌ فلان أي يُسامرُه (الفرهيدي، ١٩٨٠، ج٧، ص٢٥٥) وقد أدخل الإحصاء الآلي تكرارات الفعل "أسمر" مع تكرارات مصطلح اسم اللون الأساسي "أسمر". وهنا لا غنى عن المراجعة البشرية اليدوية لجميع سياقات تكرارات المفردة.

○ وقد أشار المرحوم محمد توفيق وردى إلى أن هناك ثلاثين نوعاً من الدبكات الكردية ...

في هذا المثال يكشف السياق أن كلمة "وَرْدِيَّ" وردت اسماً لعلم على إنسان، وليست مصطلح لون أساسي. ونظراً لتطابق بنية الكلمة في الداليتين نجد أن الإحصاء الآلي يجمع تكرارات اسم العلم على إنسان مع تكرارات مصطلح اللون الأساسي. وهنا تصبح المراجعة البشرية اليدوية لجميع سياقات تكرارات المفردة عملية مهمة.

○ ... ثنا أحمد بن إبراهيم الدورقي حدثني أسود بن سالم ثنا ...

في هذا المثال يكشف السياق أن "أسود" هي اسم علم على إنسان، ولا علاقة لها بمصطلح اللون الأساسي "أسود". ونظراً لتشابه المفردتين نطقاً

وتدويناً أدخل الإحصاء الآلي تكرارات اسم العلم على إنسان، مع تكرارات مصطلح اللون. وتصويب هذا الخطأ يتطلب المراجعة البشرية اليدوية.

○ ... ودفن في مكان يسمى (تل أسود)

في هذا المثال يكشف السياق أن "أسود" هي اسم علم على مكان، ولا علاقة لها بمصطلح اللون الأساسي "أسود". ونظرًا لتشابه المفردتين نطقًا وتدويناً أدخل الإحصاء الآلي تكرارات اسم العلم على مكان، مع تكرارات مصطلح اللون. وتصويب هذا الخطأ يتطلب المراجعة البشرية اليدوية.

○ ... بعث إليّ يوم خيبر، وأنا أرمد العين؛ فقلت: يا

رسول الله

○ ... من العين الأرمد: الذي يشتكي وجع العين

○ ... فأتيت علياً، فجنّت به أقوده، وهو أرمد

○ ... كان قد خرج من مصر -وهو أرمد- يسير في

المحفة.

○ وأن الصحيح ينظر في عين الأرمد فيرمد.

○ ... وهو أرمد شديد الرمد.

في هذا المثال يكشف السياق أن "أرمد" هي اسم علم على مرض، ولا علاقة لها بمصطلح اللون الأساسي "أرمد". ونظرًا لتشابه المفردتين نطقًا وتدويناً أدخل الإحصاء الآلي تكرارات اسم العلم على مرض، مع تكرارات مصطلح اللون. وتصويب هذا الخطأ يتطلب المراجعة البشرية اليدوية.

○ يقال: أرمد القوم. هلكت ماشيتهم وبه سمي عام الرمادة.

في هذا المثال يكشف السياق أن "أرمد" هي فعل يدل على شدة الجفاف والقحط، ولا علاقة لها بمصطلح اللون الأساسي "أرمد". ونظرًا لتشابه المفردتين نطقًا وتدويناً أدخل الإحصاء الآلي تكرارات الفعل، مع تكرارات مصطلح اللون. وتصويب هذا الخطأ يتطلب المراجعة البشرية اليدوية.

○ ... رجلاً من يهود خيبر، يعرف بسمير بن أدكن، قال في

يصول أبو حفص علينا بدرة

في هذا المثال يكشف السياق أن "أدكن" هي اسم علم على إنسان، ولا علاقة لها بمصطلح اللون الأساسي "أدكن". ونظرًا لتشابه المفردتين نطقًا وتدويناً أدخل الإحصاء الآلي تكرارات اسم العلم على إنسان، مع تكرارات مصطلح اللون. وتصويب هذا الخطأ يتطلب المراجعة البشرية اليدوية.

○ ... مع مراعاة أن بعض أنواع الزيوت تكون أدكن

لونهاً من أنواع الزيوت الأخرى

○ ... خالطته الشوائب فإن لونه يتغير إلى ما هو أدكن

(بيج)

- ... بحركة سريعة مزج لوناً أزرق، أدكن، وفرشه في منتصف اللوحة
- ... وراح يطلي جزءها العلوي بلون آخر أدكن مائل إلى الخضرة وحدده بلون أكثر غموضاً
- ... فهناك مناطق أو بقع لونها أدكن من لون بقية الجهة.
- ... اختاري ظلاً أدكن قليلاً من لون البشرة للحيلولة دون ظهور اللمعة.
- ... لأنها تجعل البشرة تبدو أدكن.
- ... امزجوا اللون الذي اخترتموه مع لون أدكن منه؛ ليعطي للوجه شكلاً بيضاً.
- استعملي ظلاً واحداً أدكن من لون الشعر الذي تتوين صباغته.
- لمزج لون أحمر الشفاه مع ظل أدكن قليلاً.
- في كل الأمثلة السابقة يكشف السياق أن المقصود هو "صيغة التفضيل" من الفعل "دكن" وليس مصطلح اسم اللون الأساسي "أدكن". وهذه الأمثلة وردت بكثرة لافتة في المدونة النصية، ونظراً لتشابه المفردتين نطقاً وتدويناً أدخل الإحصاء الآلي تكرارات صيغة التفضيل، مع تكرارات مصطلح اللون. وتصويب هذا الخطأ يتطلب المراجعة البشرية اليدوية.
- ... وطاقمه الفني المساعد يعد أحد دوافع المدرب الأرجواني لمغادرة بلاده.
- ... واصل فريق النصر تمارينه اليومية تحت إشراف الأرجواني جورج ديسلفا.
- في المثالين السابقين نجد أن السياق يوضح أن المقصود بالأرجواني هو النسب لدولة أجنبية هي دولة أوروغواي، وليس له أية علاقة بمصطلح اسم اللون الأساسي "أرجواني". ونظراً لتشابه المفردتين نطقاً وتدويناً أدخل الإحصاء الآلي تكرارات الاسم المنسوب، مع تكرارات مصطلح اللون. وتصويب هذا الخطأ يتطلب المراجعة البشرية اليدوية.
- ... قد كتبناه في ترجمة أسمر بن مضر.
- في المثال السابق نجد أن السياق يوضح أن المقصود بـ"أسمر" هو اسم علم على إنسان، وليس له أية علاقة بمصطلح اسم اللون الأساسي "أسمر". ونظراً لتشابه المفردتين نطقاً وتدويناً أدخل الإحصاء الآلي تكرارات "أسمر"

اسم علم على إنسان، مع تكرارات "أسمر" مصطلح اللون. وتصويب هذا الخطأ يتطلب المراجعة البشرية اليدوية.

○ فرض حظر شامل للتجوال في مدينة الرمادي علي خلفية التفجيرين المزدوجين... .

في المثال السابق نجد أن السياق يوضح أن المقصود بـ"الرمادي" هو اسم علم على مدينة، وليس له أية علاقة بمصطلح اسم اللون الأساسي المعروف بأل "الرمادي". ونظرًا لتشابه المفردتين نطقًا وتدوينًا أدخل الإحصاء الآلي تكرارات "الرمادي" اسم علم على مدينة، مع تكرارات "الرمادي" مصطلح اللون. وتصويب هذا الخطأ يتطلب المراجعة البشرية اليدوية.

○ فاجأت مقدمة البرامج الرياضية التلفزيونية تانيا زيتوني

مهنا جمهورها اللبناني والعربي بالانتقال من التقديم الى الإخراج.

○ ... أوقفون شارك في تجنيد جمال زيتوني

○ وقال زيتوني في اتصال هاتفي أجرته معه "الشرق الأوسط":

○ هنيئا لأخينا زيتوني، وعقبى لإنتاجات لاحقة.

في الأمثلة السابقة نجد أن السياق يوضح أن المقصود بـ"زيتوني" هو اسم علم على إنسان، وليس له أية علاقة بمصطلح اسم اللون الأساسي "زيتوني". ونظرًا لتشابه المفردتين نطقًا وتدوينًا أدخل الإحصاء الآلي تكرارات "زيتوني" اسم علم على إنسان، مع تكرارات "زيتوني" مصطلح اللون. وتصويب هذا الخطأ يتطلب المراجعة البشرية اليدوية.

○ ... وإن قال: اعصر زيتوني فما عصرت فلك نصفه

لم يجز.

في المثال السابق نجد أن السياق يوضح أن المقصود بـ"زيتوني" هو اسم علم على نبات (زيتون) منسوب لياء المتكلم (ي)، وليس له أية علاقة بمصطلح اللون الأساسي "زيتوني". ونظرًا لتشابه المفردتين نطقًا وتدوينًا أدخل الإحصاء الآلي تكرارات "زيتوني" اسم علم على نبات، مع تكرارات "زيتوني" مصطلح اللون. وتصويب هذا الخطأ يتطلب المراجعة البشرية اليدوية.

○ وكذلك جهود الشيخ محمد بشير الزيتوني وهو من

علماء تونس.

في المثال السابق نجد أن السياق يوضح أن المقصود بـ"زيتوني" هو نسب لمدينة (زيتونة) التونسية، وليس له أية علاقة بمصطلح اللون الأساسي المعروف بأل "الزيتوني". ونظرًا لتشابه المفردتين نطقًا وتدوينًا أدخل

الإحصاء الآلي تكرارات "زيتوني" نسبة لمدينة زيتونة التونسية، مع تكرارات "زيتوني" مصطلح اللون. وتصويب هذا الخطأ يتطلب المراجعة البشرية اليدوية.

خلاصة: يتضح من إشكاليات الإحصاء الآلي التي تبين عند إحصاء تكرارات شبوع مصطلحات الألوان الأساسية في المدونة العربية أن الوصول إلى قيم تكرارات دقيقة ومنضبطة بالاعتماد على الإحصاء الآلي فقط أمر مستبعد. وأن هذا المستوى من الدقة يتطلب البحث بالكشاف السياقي لمراجعة سياق كل مصطلح من المصطلحات (مصطلح اللون الأساسي منفردًا + المصطلح بالسوابق + المصطلح باللواحق + المصطلح بالسوابق واللواحق).

تكمن المشقة في الفحص اليدوي لسياقات المدونات المليارية -مثل المدونة العربية- أننا نعالج عادة أرقامًا كبيرة جدًا. وفي دراستنا الحالية تطلب الفحص اليدوي لسياقات ستة عشر مصطلحًا لأسماء الألوان الأساسية فحص ستمئة وخمسة عشر ألفًا وسبعمئة وخمسة وسبعين (٦١٥٧٥٧) تكرارًا، وكل تكرار له نصه، وسياقه المستقل.

أما عن مراجعة هذه التكرارات في سياقها فقد تم عن طريق البحث في الكشاف السياقي عن كل مصطلح، ومراجعة سياق كل تكرار، ومن ثم حذف التكرارات التي يكشف سياقها أنها لا تدخل في مصطلحات اللون الأساسية. ومن ثم تم إعادة تصويب قيم التكرارات. وعند كتابة معجم الشبوع اللغوي لمصطلحات الألوان الأساسية في صورته النهائية تم ذكر التكرارات بعد المراجعة والتدقيق وتصويب الأرقام. ولهذا نجد في الصورة النهائية للمعجم تعديلات واضحة في تكرارات جميع مصطلحات الألوان الأساسية في العربية.

٩- الصورة النهائية للمعجم.

١/٩- المكونات التعريفية للمعجم.

حرصنا عند تأليف معجم الشبوع اللغوي لمصطلحات الألوان الأساسية في اللغة العربية على أن نصل المكونات التعريفية للمعجم، توضيحًا للمُطَّلِع، وتيسيرًا إجرائيًا لحوسبة هذا المعجم فيما بعد. ووفقًا لتصنيف المكونات التعريفية إلى مكون سائد، ومكون شائع، ومكون مائز، ومكون تكميلي (عبد الصبور، ٢٠١٦، ص ١٧٩)، فقد شمل هذا المعجم مكونات تعريفية سائدة في معجم الشبوع، وهي المدخل، والتعريف، ومجموع تكرارات المصطلح. وأضافنا له مكونات تعريفية مائزة، وهي صورة اللون الشائعة عنه، وأكثر صور المصطلح استعمالاً، ونصوص الاستعمال، ومكونًا تكميليًا هو ترجمات

المصطلح. وعلى هذا كانت المكونات التعريفية لهذا المعجم كما يلي على الترتيب:

- ١ **المدخل:** في المدخل اعتمدنا جذع المصطلح (المفرد المذكر النكرة للمصطلح). وسيتم تمييز مصطلحات الألوان الأساسية التي لا خلاف حولها بعلامة النجمة بين قوسين (*) بعد المصطلح، باعتبارها مصطلحات ألوان أساسية ذات مرتبة أولى. ولم نميز غيرها من المصطلحات التي يوجد ملاحظة حول كونها مصطلح لون أساسي في العربية على اعتبارها مصطلحات ألوان أساسية ذات مرتبة ثانوية.
- ٢ **صورة اللون:** لأن اللون يدرك بالنظر، أضفنا صورة تمثل الشائع في هذا اللون الأساسي.
- ٣ **التعريف:** ركزنا في التعريف على ذكر مثال واضح للون ما تيسر ذلك، مع التعريف بالمنظومة المصطلحية له. واعتمدنا في التعريف على مجموعة من المعاجم العربية القديمة والحديثة، وقدمناها على المعاجم الأجنبية؛ نظراً لكوننا نبحث عن المفهوم العربي لمصطلحات الألوان الأساسية.
- ٤ **مجموع تكرارات المصطلح:** ذكرنا قيمة التكرار النهائية لشيوع المصطلح في المدونة العربية بعد المراجعة السياقية.
- ٥ **أكثر صور المصطلح استعمالاً:** ذكرنا الصورة الأكثر شيوعاً من صور استعمال المصطلح، مع ذكر تكرار هذه الصورة بعد المراجعة السياقية.
- ٦ **نصوص الاستعمال:** قدمنا لكل مصطلح عشرة نصوص للاستعمال الفعلي مما ورد في المدونة العربية. وراعينا عند انتقائها أن تبتعد النصوص -قدر الإمكان- عن الغموض، واللبس. كما قمنا بتصحيح ما ورد في النص الأصلي من أخطاء إملائية أو إعرابية منعاً لوقوع القارئ في لبس.
- ٧ **ترجمات المصطلح:** للتوضيح للقارئ الأجنبي المتعلم للغة العربية قدمنا ترجمات لمصطلح اللون الأساسي العربي باللغات الإنجليزية، والفرنسية، والألمانية.

معجم الشبوع اللغوي لمصطلحات الألوان الأساسية في اللغة العربية

أَحْمَر (*)

التعريف: ما لونه كلون الدَّم (عمر، 2008، ج1، ص 557).
وأحمر جمعه حُمْر. والمؤنث حُمْرَاء، ويجمع المؤنث على
حُمْراوات وحُمْر. (الفيومي، 1987، ج1، ص150)
مجموع التكرارات: 67964 مرة
أكثر صورته استعمالاً: الأَحْمَر (49210مرة)

نصوص الاستعمال:

- ١ تلبس لباساً حريزاً أحمر.
- ٢ حبسوه في علب مغلقة بالشمع الأحمر.
- ٣ حصل لابعوه على كارت أحمر خلال الدور الأول.
- ٤ حضرت الفرق الطبية من الهلال الأحمر.
- ٥ في الوسط باقة من الورد الأحمر في حب مصر.
- ٦ قال الصليب الأحمر: إن العشرات قتلوا.
- ٧ تجهز نصف كيلو لحم أحمر مفروم.
- ٨ نحضر كمية بسيطة من ورق البصل الأحمر.
- ٩ هناك دول مطلة على البحر الأحمر من بينها مصر.
- ١٠ وحدتنا الوطنية خط أحمر يحب الوقوف عنده.

الترجمات:

English	Français	Deutsch
Red	Rouge	Rot

التعريف: لُون كلون الفَحْم ينتج من امتصاص أشعة النُور
امتصاصاً تاماً. ضده أبيض. ويجمع أسود على سُود وسُودان،
ومؤنثه سَوْدَاء، ويجمع المؤنث على سَوْداوات وسُود. (عمر،
2008، ج2، ص1130)
مجموع التكرارات: 64104 مرة
أكثر صورته استعمالاً: الأسود (40445مرة)

نصوص الاستعمال:

١. أكد المنتخب الأوزبكي خلال الجولتين الماضيتين أنه
الحصان الأسود لكرة القدم الآسيوية.
٢. إلى جانب تفريغ محتويات الصندوق الأسود الذي تم
انتشاله.
٣. شاعر فارس، من أغربة العرب. كان أسود اللون (أخذ
السواد من أمه ندية).
٤. صادق الباحثون على مركبات polyphenolic في الشاي
الأسود التي تساعد في كبح الشعور بالجوع.
٥. ضعي ملعقة فلفل أسود.
٦. كان اسمه محمد بن إسماعيل، وكان أسود.
٧. كان عيِّداً أسود، من موالى بني أسد.

أَسْوَد (*)

٨. كل الجميع يوم أكل الثور الأسود.
٩. نساؤنا يلبسن ثوباً أسود، وتحتة ثوب آخر، فلا جناح عليهن.
١٠. يقع الحجر الأسود في الزاوية الشرقية من الكعبة المشرفة.

الترجمات:

English	Français	Deutsch
Black	Noir	Schwarz

التعريف: ضدُّ الأسود، من البَيَاضِ. (الزبيدي، 1979، ج18، ص250)، وهو ما كان بلون الثلج النقي، أو ملح الطعام النقي. (عمر، 2008، ج1، ص270) وجمع أبيض ببيض، والمؤنث منه ببيضاء، ويجمع المؤنث على ببيض وببيضاوات.

مجموع التكرارات: 60021 مرة

أكثر صوره استعمالاً: الأبييض (39786 مرة)

نصوص الاستعمال:

- ١ ارتفاع مياه البحر الأبيض المتوسط.
- ٢ تترقب جماهير الأبيض المباراة المصرية التي تجمع بين الزمالك والإفريقي التونسي.
- ٣ سبق اتهامه في قضية سلاح أبيض.
- ٤ شامخة البناء وأسوارها من الرخام الأبيض
- ٥ في الغالب يقدرون الأمور باللونين الأبيض والأسود.
- ٦ كلها عوامل تعجل بظهور الشعر الأبيض.
- ٧ الكوثر نهر أعطانيه الله في الجنة، ترابه مسك، أبيض من اللبن.
- ٨ مثل رؤوس الكعك مرشوش بسكر أبيض.
- ٩ مرشح بارز ينسحب من سباق الحزب الجمهوري إلى البيت الأبيض.
- ١٠ يتصاعد اللهب، وسرعان ما ينتشر دخان أبيض.

الترجمات:

English	Français	Deutsch
White	Blanc	Weiss

التعريف: لون الحشائش الغضة. (عمر، 2008، ج1، ص655) وجمعه خُضْرٌ، ومؤنثه خُضْرَاءُ، ويجمع المؤنث على خُضْرٌ، وخُضْرَاوَات. (عمر، 2008، ج1، ص656)

مجموع التكرارات: 37192 مرة

أكثر صوره استعمالاً: الأخضر (26315 مرة)

نصوص الاستعمال:

- ١ اختاري تفاعاً لونه أخضر وحجمه صغير ورائحته طيبة.
- ٢ تعطي للمستثمر العالمي الضوء الأخضر للعمل.
- ٣ سيواصل اللاعب العطاء على المستطيل الأخضر.

أبيض (*)

أخضر (*)

معاجم الشبوع والمدونات اللغوية العربية
دراسة على ضوء لسانيات المدونات

- ٤ الشاي الأخضر يحمي الجلد.
٥ عليه ثوب أخضر، والهواء يلعب بشيئته.
٦ كانا يرتديان القميص الأخضر الخاص بالأطباء
والمرضى اللبيين.
٧ لم يجروا الفنان على استخدام اللون الأخضر.
٨ لم يعد هناك عود أخضر في الحقول.
٩ لو اشتعلت نار الحرب لأحرق كل أخضر ويايس.
١٠ منح 5 آلاف تصريح عمل داخل الخط الأخضر لعمال
بناء فلسطينيين.

الترجمات:

English	Français	Deutsch
Green	Vert	Grün

التعريف: ما كان بلون السماء الصافية. (عمر، 2008، ج2، ص 981) والفعل منه: زرق زرقاً، فهو أزرق، وهي زرقاء. (الزبيدي، 1979، ج25، ص394)
مجموع التكرارات: 21275 مرة
أكثر صورته استعمالاً: الأزرق (14876 مرة)
نصوص الاستعمال:

1. ... بخصوص نسب اللاعبين من أصول أفريقية وعربية داخل المنتخب الأزرق والتي هزت كرة القدم الفرنسية في الآونة الأخيرة.
2. التغير المناخي الذي يشهده العالمي يهدد الكوكب الأزرق.
3. شيء من موسيقى الدانوب الأزرق.
4. عثر النقابون على شققة كبيرة من إناء مصنوع من الحجر الأزرق.
5. عضهم ثعبان التعصب بنابه الأزرق.
6. فكأنها فيه بساط أزرق ... وكأنه فيها طراز مذهب.
7. كان أشقر اللحية، أزرق العينين، عليه وقار وهيبة.
8. كان سلفاكير في زيارة لمنطقة كردوفان والنيل الأزرق.
9. ليس عليه سوى قميص أزرق يلبسه صيفاً وشتاء.
10. يضيف أن تحول لون اللسان إلي الأزرق دليل على الإصابة بهبوط القلب.

الترجمات:

English	Français	Deutsch
Blue	Bleu	Blau

التعريف: ما لونه كلون الذهب. يجمع على صُفر، ومؤنثه صُفراء، ويجمع المؤنث على صُفراوات وصُفر. (عمر، 2008، ج2، ص1301)
مجموع التكرارات: 13755 مرة
أكثر صورته استعمالاً: الأصفر (7817 مرة)

أزرق (*)



أصفر (*)

نصوص الاستعمال:

١. ترك البنت الجميلة البيضاء ذات الشعر الأصفر ترجع لوحدها.
٢. تسكنه رائحة القمح الأصفر.
٣. حكم المباراة علاء عبد القادر استخدم الكارت الأصفر أربع مرات قسمت بالتساوي بين الفريقين.
٤. قد صار لوني أصفر كأنه الزعفران.
٥. القلادات والسلاسل مصنوعة من الذهب الأصفر.
٦. نسخ من جرائد ورقها أصفر.
٧. هو صاحب الطبل المستقر عندهم من نحاس أصفر كان معه في فتح عكة يضربون به.
٨. وضع القاضي على رأسه منزرراً أصفر.
٩. يعطي الكشف ضوءاً أصفر تحذيراً.
١٠. يمكنك إضافة قليل من الملح والكرم الأصفر.

الترجمات:

English	Français	Deutsch
Yellow	Jaune	Gelb

التعريف: ما كان بلون الرّماد، وهو لُون بين الأسود والأبيض (عمر، 2008، ج2، ص941)، أو غيرة تضرب إلى البياض. ويطلق عليه -أيضاً- أربد، رصاصي، بوشر، أشهب. (دوزي، 1997، ج5، ص213) وجمعه (رَمَادِيُون ورَمَادِيِين)، ومؤنثه رَمَادِيَّة، ويجمع على رَمَادِيَّات.

مجموع التكرارات: 11512 مرة

أكثر صوره استعمالاً: الرّماديّ (6549 مرة)

نصوص الاستعمال:

١. إن العلاقات الانسانية تحولت إلي اللون الرمادي الكريه.
٢. تتلون السحب بلون رمادي يجلب الحزن والتعاسة للناس.
٣. تصاعد عامود من الدخان الرمادي.
٤. عليه فائلة خفيفة من الصوف الرمادي لها ياقة وفتحة قصيرة بأزرار مغلقة.
٥. كما يعطي اللون الرمادي بتدرجاته الخفيفة إحساساً بالحياة.
٦. متى يتحول شعر الشخص إلى اللون الرمادي؟
٧. هو محصن بكتل من المكعبات الحجرية ذات اللون الرمادي القائم.
٨. يحتوي علي مسحوق رمادي يشتهبه في كونه باروداً.
٩. يرتدي قميصاً قطنياً رمادي اللون.
١٠. ينصحون باستعمال اللون الرمادي في ظلال العيون بدلاً من اللون الأسود.

رَمَادِيّ (*)

معاجم الشبوع والمدونات اللغوية العربية
دراسة على ضوء لسانيات المدونات

الترجمات:			
English Gray	Français Gris	Deutsch Grau	
<p>التعريف: لون الورد (دوزي، 1997، ج11، ص55) وهو لون أحمر باهت بين الأحمر والأبيض. وجمعه (وَرْدِيُون، وَوَرْدِيِين)، ومؤنثه وَرْدِيَّة ويجمع على وَرْدِيَّات.</p> <p>مجموع التكرارات: 5767 مرة</p> <p>أكثر صورته استعمالاً: الْوَرْدِيَّ (3363مرة)</p> <p>نصوص الاستعمال:</p> <ol style="list-style-type: none"> 1. أفواج من المحترفين الصينيين مهمتهم انتقاء أفضل، وأندر مقاطع الجرائيت الوردية، والرخام النادر. 2. إن الجولة في العاصمة بوينس إيرس تبدأ من أمام القصر الوردية. 3. تكون الخياشيم أثناء فتحها ذات لون أحمر ووردي، وخالية من المخاط. 4. تم عمل تبييط للأرصفة، ودهان الأرصفة بالكامل باللون الوردية. 5. سرعان ما أفيق بعد هنيهة وجيزة من حلمي الوردية الجميل. 6. كانت الرسالة ملفوفة بشريط ووردي، وتفوح عطراً. 7. له رؤية استشرافية لمستقبل ووردي للحياة. 8. المخرج السينمائي شون ليفي في النسخة الجديدة لفيلم (الفهد الوردية) الذي أخرجه بلاك إدوارد في عام 1963م. 9. وكانت تخضب كفيها، وقدميها بالحناء ذات اللون الوردية الأحمر. 10. يذكر أن اللون الوردية يعد الرمز العالمي لحملات التوعية من أخطار سرطان الثدي. 			<p>وَرْدِيَّ</p> 
الترجمات:			
English Pink / Rose	Français Rose	Deutsch Rosa / Pink	
<p>التعريف: لُون بَيِّنُ النَّيَّاضِ وَالسَّوَادِ. (دوزي، 1997، ج6، ص143) فَهُوَ لَوْنٌ يَضْرِبُ إِلَى سَوَادٍ خَفِيٍّ. (ابن منظور، 1993، ج4، ص376) يَكُونُ فِي أَلْوَانِ النَّاسِ، وَغَيْرِهَا. (الزبيدي، 1979، ج12، ص71) وَهُوَ مَنْ كَانَ شَعْرُهُ أَسْوَدَ، وَلَوْنُهُ أَسْمَرَ. (دوزي، 1997، ج6، ص143). وَفَعْلُهُ سَمَّرَ بِالضَّمِّ، فَهُوَ أَسْمَرٌ، وَالْأُنثَى سَمْرَاءُ. (الفيومي، 1987، ج1، ص288)</p> <p>مجموع التكرارات: 4857 مرة</p> <p>أكثر صورته استعمالاً: الْأَسْمَرُ (1760مرة)</p> <p>نصوص الاستعمال:</p> <ol style="list-style-type: none"> 1. أحسست بوجود مشادة بين الوزيرة وإحدى الشابات 			<p>أَسْمَر</p> 

١. لونها أسمر، عرفت بعد ذلك أنها من العاملين في الأهرام.
٢. الأغذية التي تشعر الجسم بالشبع كالبقوليات والمواد النشوية كالخبز الأسمر.
٣. اكتشف العنديلِبَ الأسمر، عبد الحليم شبانة، الإذاعي الكبيرُ حافظ عبد الوهاب.
٤. عهدي به أسمر، أسود اللحية والرأس.
٥. في أوقات أخري نجد لون الرغيف أسمر كأنه مخلوط بتراب.
٦. قد كان الإمام أحمد رجلاً رقيقاً أسمر اللون كثير التواضع.
٧. كان أحمد عيد يناديه في الفيلم باسم بكار بسبب لونه الأسمر.
٨. كان الزبير -رضي الله عنه- أسمر، ربعة، معتدل اللحم خفيف اللحية.
٩. كان رجلاً أسمر اللون، طويل القامة.
١٠. لا تضرب كلبي الأسمر الصغير.

الترجمات:

English	Français	Deutsch
Light brown skin	Peau brun clair	Hellbraune Haut

التعريف: ما كان بلون قشرة البرتقال الناضج. (عمر، 2008، ج1، ص181) وجمعه (بُرْتَقَالِيُون وِبُرْتَقَالِيِين)، والمؤنث منه بُرْتَقَالِيَّة. ويجمع على بُرْتَقَالِيَّات.

مجموع التكرارات: 4291 مرة

أكثر صوره استعمالاً: البُرْتَقَالِيَّ (2373 مرة)

نصوص الاستعمال:

١. اكتشفت أنني إنما أتأمر على عصفور صغير جميل برتقالي الصدر، يدعى أبو الحناء.
٢. انتهت الدرب بجنة واسعة من أزهار عباد الشمس البرتقالي.
٣. بعد وضع محلول الأيودين برتقالي اللون، ثم النشاء يتحول اللون إلى الكحلي.
٤. تعكر الجو، وتغير إلى اللون البرتقالي، بسبب العواصف الترابية الشديدة.
٥. جلست على مصطبة خشبية ملتفة بشال برتقالي اللون.
٦. الكاروتينويدات هي التي تعطي اليوسفي لونه البرتقالي.
٧. كان يرتدي معطفاً برتقالي اللون.
٨. من المؤكد أن وجود شيء ما بلون برتقالي في المنزل من شأنه أن يجعلك تبتسم، وتفرح.
٩. يتبدد وهج المغيب البرتقالي على صفحة الماء.

بُرْتَقَالِي



معاجم الشبوع والمدونات اللغوية العربية
دراسة على ضوء لسانيات المدونات

١٠. يوجد نوع من الفل، ولكنه برتقالي اللون، وهو ما يسمى بالنرجس.

الترجمات:

English	Français	Deutsch
Orange	Orange	Orange

التعريف: ما كان لونه لونَ زهرة البَنَفْسَج، أزرَقَ مائل للحمرة. (عمر، 2008، ج1، ص249) وجمعه (بَنَفْسَجِيُون وبَنَفْسَجِيِين)، ومؤنثه بَنَفْسَجِيَّة، ويجمع على بَنَفْسَجِيَّات.

مجموع التكرارات: 4228 مرة

أكثر صورهِ استعمالاً: البَنَفْسَجِيَّة (2301 مرة)

نصوص الاستعمال:

١. إضافة إلى اللون البنفسجي الذي يعتبر موضة الأيام القادمة.
٢. إن عملية التبييض بالأشعة فوق البنفسجية قد تسبب الحساسية عند بعض المرضى.
٣. أول ظهور لياسر القحطاني بالقميص البنفسجي.
٤. طارت فراشتي بجناح بنفسجي إلى صحرائك.
٥. كما تعطي حماية من الأشعة فوق البنفسجية الضارة.
٦. لا بد أن يكون هناك التوت بلونه البنفسجي الغامق.
٧. اللحوم المستوردة يتم ختمها بختم بنفسجي اللون.
٨. لم تجد لوحة السوسنات البنفسجية لمونيه أيضاً من يشتريها.
٩. النور الخارجي يخترق ستائر الغرفة عبر النافذة بلون بنفسجي خامل ورقيق.
١٠. هؤلاء هم الفريسيون بملابسهم البنفسجية الواسعة ذات الأطراف المنسوجة بالذهب.

الترجمات:

English	Français	Deutsch
Purple	Violet	Lila

التعريف: لون البُنِّ المطحون. (مجمع اللغة العربية، 2004، ج1، ص72) وهو لون الشوكولاتة. وجمعه (بُنِّيُون وبُنِّيِين)، ومؤنثه بُنِّيَّة، وجمعها بُنِّيَّات.

مجموع التكرارات: 11602 مرة

أكثر صورهِ استعمالاً: البُنِّيَّة (1533 مرة)

نصوص الاستعمال:

١. استخدام أغطية للفراش ذات ألوان هادئة مثل اللون البني لتوفير الهدوء للسجناء.
٢. إن نبات المر عبارة عن مادة صمغية ذات لون بني مائل للاحمرار.
٣. جلبت أحجاره المسطحة ذات اللون البني الغامق من سفح الجبل الجنوبي.

بَنَفْسَجِي



بُنِّي (*)



٤. ذلك بملء الفراغات فيهما باللون البني.
٥. الربع الخالي أرض جرداء يغلب عليها لون بني ضارب إلى الحمرة.
٦. ظهور بقع يتراوح لونها من اللون البني الفاتح الى اللون البني الغامق على مناطق الجبهة.
٧. عبارة عن جملٍ صغير منحوت من الخشب ذو لون بني وملمس شمعي ناعم.
٨. لون البول داكن يميل إلى اللون البني.
٩. الميلانين يعطي الجلد اللون البني.
١٠. يجعل الشعر البني الباهت أكثر بريقاً وإشراقاً.

الترجمات:

English	Français	Deutsch
Brown	Brun	Braun

التعريف: لَوْنٌ أَحْمَرٌ غَامِقٌ يَمِيلُ إِلَى اللَّوْنِ الْبَيْضِجِيِّ. (دوزي، 1997، ج8، ص 55) دَرَجَةٌ لَوْنِيَّةٌ تَفَعُّ بَيْنَ الْبَيْضِجِيِّ وَالْأَحْمَرِ. (عمر، 2008، ج1، ص82) ويطلق عليه أيضا أَرْجَوَانٌ، وَفَرْفِيرِيَّةٌ (دوزي، 1997، ج8، ص 55) وَقَطِيفِيَّةٌ. (دوزي، 1997، ج4، ص 459).

مجموع التكرارات: 869 مرة

أكثر صورته استعمالاً: الأَرْجَوَانِيَّةُ (392 مرة)

نصوص الاستعمال:

١. الأنوف قد اكتست بلون أرجواني.
٢. الثمار التفاحية تكون أنصال أوراقها ذات لون أرجواني.
٣. شبه فراشة غريبة بها ظلال من اللونين الأرجواني الفاتح والأحمر.
٤. الطحال عضو إسفنجي رقيق أرجواني اللون يقع خلف المعدة وإلى يسارها.
٥. عيون أطفالها التي تشبه صدف الفونيق الأرجواني.
٦. غربت الشمس وراء الأفق، واصطبغت السماء بلون أرجواني.
٧. قد ارتدى العباءة ذات اللون الأرجواني موشاة بغزارة بالذهب.
٨. كانت ترتدي ثوبا بلون أرجواني فاتح
٩. مزج قطرتين من زيت المرمرية الأرجواني ليمنحك إحساساً بالتفاؤل والمرح.
١٠. يمكنك ارتداء الأرجواني مع الجينز الرمادي لأناقة مبهرة.

الترجمات:

English	Français	Deutsch
Amaranthine	Amarante	Amaranth

أَرْجَوَانِيَّةٌ



زَيْتُونِيّ

التعريف: ما كان بلون الزَيْتُون، أخضر مائل للصفرة. (عمر، 2008، ج2، ص1013). جمعه (زَيْتُونِيُون، وَزَيْتُونِيِين)، ومؤنثه زَيْتُونِيَّة، ويجمع على زَيْتُونِيَّات.

مجموع التكرارات: 190 مرة

أكثر صوره استعمالاً: الزَيْتُونِيّ (98مرة)

نصوص الاستعمال:

١. أربعة أشخاص كانوا يرتدون البدلات ذات اللون الزيتوني.
٢. أطلق على النصب المبني من الجبس الأبيض المصبوغ باللون الزيتوني اسم "نصب الناجين".
٣. أنصحك باستخدام اللون الزيتوني.
٤. انطلقت بعض طائرات الهليكوبتر العسكرية بلونها الزيتوني.
٥. أوراقها بسيطة معنقة سهمية متقابلة ذات لون أخضر داكن (زيتوني) تخرج من أباطها البراعم الزهرية.
٦. جمال لون عينيها الأزرق المائل إلى الزيتوني.
٧. حيث استخدم اللون الذهبي إلى جانب الأرجواني الداكن والأخضر الزيتوني.
٨. لم أجد سوى سبعة ملفات ذات لون أخضر زيتوني من ورق مقوى.
٩. محبرة تزدان بالألوان ما بين أحمر داكن، وكحلي، وأخضر زيتوني، وبني إلى الأسود.
١٠. وراء مقعد السيارة الخلفي صندوق مستطيل الشكل زيتوني اللون.

الترجمات:

English	Français	Deutsch
Olive color	Couleur Olive	Olivfarbene Farbe

أَدْكَن

التعريف: لَوْنُ الأَدْكَنِ كَلَوْنِ الحَزِّ الَّذِي يَضْرِبُ إلى العُبرة بَيْنَ الحُمْرَةِ وَالسَّوَادِ (ابن منظور، 1993، ج13، ص157) وَالْفِعْلُ دَكَنَ يَدْكُنُ دَكْنًا. (أبو منصور، 2001، ج10، ص73). والمؤنث دكناء. (ابن سيده، 1996، ج1، ص204)

مجموع التكرارات: 102 مرة

أكثر صوره استعمالاً: أَدْكَن (62مرة)

نصوص الاستعمال:

١. إن الغمام الأبيض محمود، والأسود مزموم، والأدكن بينهما في الرداءة والجودة.
٢. برز بوضوح على أرضية من الحصى الأدكن.
٣. دادبي: هي حب مثل الشعير إلى حمرة ما، وزهره أطول، وأدق، أدكن، مر.
٤. رأيت أنس بن مالك عليه جبة خز دكناء، ومطرف خز

أدكن.		
٥	الشوارع قد بنيت بالطوب الأدكن.	
٦	طائر على هيئة القمري جلبابه أدكن، وعيناه، ومنقاره حمر، وجناحاه مخططان بالسواد.	
٧	قد كساها ريشها الأدكن.	
٨	كأنه والغيم ثوب أدكن.	
٩	نظرت إلى البدر عند الخسوف ... وقد شين منظره الأزين // كما سفرت صفحة للحبيب ... يحجبها برقع أدكن.	
١٠	هذه العجرية تخطر بثوبها الهفهاف، تلاحق شاباً أدكن اللون.	
الترجمات:		
English	Français	Deutsch
Black	bean	Schwarze
color	Couleur	Bohnenfarbe
	haricot noir	
<p>التعريف: الْأَدِي عَلَى لَوْنِ الرَّمَادِ، وَهُوَ غُبْرَةٌ فِيهَا كُدْرَةٌ؛ وَمِنْهُ قِيلَ لِلْبُعُوضِ رُمْدٌ. (ابن منظور، 1993، ج3، ص185) وَنَعَامَةٌ رَمْدَاءٌ فِيهَا سَوَادٌ مُنْكَشِفٌ كَلَوْنِ الرَّمَادِ. (ابن سيده، 2000، ج9، ص330) والماء "الرمد" أي الذي صار على لون الرماد. (الكجراتي، 1967، ج2، ص377)</p> <p>مجموع التكرارات: 2. (مرتان)</p> <p>أكثر صورته استعمالاً: الأُرْجَوَانِيّ 2. (مرتان)</p> <p>نصوص الاستعمال:</p> <p>١. أخضر مدهام، وأورق خطباني، وأرمد دراني.</p> <p>٢. ثياب كأنها القراطيس، وشطر عليهم ثياب رمد جمع أرمد أي أكر، فدخل البيت المعمور ودخل معه.</p>		
الترجمات:		
English	Français	Deutsch
Ash color	Couleur cendre	Aschefarbe

أرمد



10- الخاتمة وأهم النتائج.

ناقشت هذه الدراسة عدداً من التساؤلات، هي: إلى أي مدى يمكن الاعتماد على المدونات العربية الحديثة في إنشاء معاجم الشبوع؟ وأي المدونات اللغوية العربية الحديثة أكثر فعالية في إحصاءات شبوع المفردات؟ وهل يمكن الاعتماد كلياً على الإحصاء الآلي الموجود في المدونات؟ وما إنشكاليات إحصاء الشبوع عند صناعة المعجم؟ وما المحتوى التعريفي المناسب لمعجم الشبوع؟

كما سعت هذه الدراسة لتحقيق عددٍ من الأهداف النظرية، وهي: إلقاء الضوء على أهمية معاجم الشبوع اللغوي لمفردات اللغة العربية. ومعرفة إلى أي مدى يمكن الاعتماد على المدونات العربية الحديثة في صناعة معاجم الشبوع اللغوي. واستيضاح أكثر المدونات العربية فعالية في الإحصاء الآلي للمفردات. وتحديد إشكاليات إحصاء الشبوع آلياً في المدونات العربية الكبرى. واكتشاف المحتوى التعريفي الأمثل لمعجم الشبوع. بالإضافة إلى هدفين تطبيقيين، هما: قياس شبوع مصطلحات الألوان الأساسية في اللغة العربية. وإعداد معجم شبوع لمصطلحات الألوان الأساسية في اللغة العربية. وقد قدمت هذه الدراسة عدة نتائج عامة، ومنها:

- ألفت الدراسة الضوء على أهمية معاجم الشبوع اللغوي.
 - وضحت الدراسة أنه لا يمكن الاعتماد كلياً على الباحث الآلي في المدونات العربية في صناعة معاجم الشبوع اللغوي. والحصول على نتائج دقيقة يتطلب الفحص اليدوي للسياقات من المختصين.
 - أثبتت الدراسة أن جودة الإحصاء الآلي ترتبط ارتباطاً قوياً بأن تكون جميع مفردات المدونة مشكولة شكلاً كلياً بدقة عالية.
 - وضحت الدراسة أن ثمة إشكاليات متنوعة في الإحصاء الآلي لشبوع المفردات في المدونات العربية.
 - نجحت الدراسة في تقديم إحصاء شبوع لمصطلحات الألوان الأساسية في اللغة العربية.
 - قدمت الدراسة معجم شبوع مصطلحات الألوان الأساسية في اللغة العربية.
 - تمثلت المكونات التعريفية لمعجم شبوع مصطلحات الألوان الأساسية في اللغة العربية في سبعة مكونات تعريفية هي:
 - المدخل.
 - صورة اللون.
 - التعريف.
 - مجموع تكرارات المصطلح.
 - أكثر صور المصطلح استعمالاً.
 - نصوص الاستعمال.
 - ترجمات المصطلح.
- كذلك قدمت هذه الدراسة عددًا من النتائج التفصيلية، ومنها:

- مثلت قوائم شيوع الكلمات البداية المبكرة لمعاجم الشيوخ اللغوي. وقد عرفت العربية مجموعة من قوائم شيوع الكلمات من أشهرها قائمة مكة للمفردات الشائعة.
- فيما يتعلق بشيوع المفردات لا يجوز الاعتماد على الحدس الذاتي أو المقدرة الشخصية.
- ثمة بعض المشاريع غير المكتملة لصناعة معجم شيوع لغوي عام عربي-عربي.
- يعد معجم الشيوخ للغة العربية (A Frequency Dictionary of Arabic)، من إعداد روتليدج، هو المعجم الوحيد الذي اعتمد على مدونة نصية مليونية. وهو معجم تعليمي ثنائي اللغة.
- المدونة العربية لمدينة الملك عبد العزيز هي الأكثر ثراء في المحتوى، ودقة في البحث الآلي من بين المدونات العربية الحديثة.
- ظهر مفهوم مصطلحات الألوان الأساسية مع كتاب برنت برلين، وبول كي في عام ١٩٦٩م.
- الخصوصية الثقافية للألوان لا تمنع أن هناك خصوصية لبعض الألوان.
- ثمة بعض الشروط لا اعتبار مصطلح اللون أساسياً.
- يمكن أن نتوسع في مصطلحات الألوان الأساسية في اللغة العربية إلى ستة عشر مصطلحاً، وهي: أحمر، أسود، أبيض، أخضر، أزرق، أصفر، رمادي، وردي، أسمر، برتقالي، بنفسجي، بني، أرجواني، زيتوني، أدكن، أرمد.
- عند صناعة معجم الشيوخ اللغوي لا بد من تحديد حدود المفردة وفق ما يناسب البحث الآلي في المدونة اللغوية.
- تقدم المدونة العربية مجموعة من خيارات البحث التي تساعد على إنشاء معجم للشيوخ اللغوي، لكنها لا تستغني عن التدقيق البشري للمفردات في سياقاتها.
- ثمة تباين كبيرة بين نتائج البحث الآلي للمصطلح المشكول شكلاً كلياً، ونتائج البحث عن المصطلح ذاته بدون شكل.
- ثمة تباين كبير بين تكرارات البحث الآلي للمصطلح قبل المراجعة السياقية، والتكرارات بعد المراجعة السياقية اليدوية.

- يعد عدم ضبط بنية الكلمات بالشكل من إشكاليات البحث الآلي في المدونة العربية، والذي يمكن معالجته بالشكل الكلي لجميع مفردات المدونة.
- يعد وجود أخطاء في التدوين من إشكاليات البحث في المدونة العربية، والذي يمكن معالجته بالمراجعة والتدقيق لنصوص المدونة.
- يعد اتحاد اللفظ واختلاف المعنى من إشكاليات البحث في المدونة العربية. والتي تتطلب تدخلاً بشرياً لفهم دلالة المفردات في سياقاتها.
- تثبت نتائج الإحصاء التحليلي أنه لا يوجد ارتباط ذو دلالة بين قيم الإحصاء للمصطلح المشكول والمصطلح غير المشكول، وقيم إحصاء المصطلح قبل المراجعة السياقية، وبعد المراجعة السياقية، عند مستوى الدلالة ٠.٠٥.
- تثبت نتائج الإحصاء التحليلي أنه لا يوجد فروق ذات دلالة بين قيم الإحصاء للمصطلح المشكول والمصطلح غير المشكول، وقيم إحصاء المصطلح قبل المراجعة السياقية، وبعد المراجعة السياقية، عند مستوى الدلالة ٠.٠٥.
- مصطلح بنفسجي وكذلك صورته المنبثقة عنه هو مصطلح اللون الأساسي الوحيد الذي لا يستخدم في دلالة أخرى غير دلالته على اللون، ولا يلتبس في استعماله مع مفردات أخرى، وتكرارات شيعه قبل المراجعة السياقية لم تختلف عن تكرارات شيعه بعد المراجعة السياقية.

وختاماً، توصي هذه الدراسة بالمزيد من الدراسات حول الشبوع اللغوي لألفاظ العربية، وإضافة المزيد من معاجم الشبوع العامة، والاصطلاحية. وزيادة الاعتناء بالمدونات العربية الحاسوبية.

والله ولي التوفيق.

قائمة المصادر والمراجع:

أولاً- المراجع العربية.

- ١ ابن دريد، أبو بكر محمد بن الحسن. (١٩٨٧). *جمهرة اللغة*. تحقيق: رمزي منير بعلبكي. دار العلم للملايين - بيروت. ط ١.
- ٢ ابن سيدة، أبو الحسن علي بن إسماعيل. (١٩٩٦). *المخصص*. تحقيق: خليل إبراهيم جفال. دار إحياء التراث العربي - بيروت، ط ١.
- ٣ ابن سيدة، أبو الحسن علي بن إسماعيل. (٢٠٠٠). *المحكم والمحيط الأعظم*. تحقيق: عبد الحميد هنداوي، دار الكتب العلمية، بيروت، ط ١.
- ٤ ابن منظور، محمد بن مكرم بن علي أبو الفضل جمال الدين. (١٩٩٣). *لسان العرب*. دار صادر، بيروت، ط ٣.
- ٥ أبو منصور، محمد بن أحمد بن الأزهر الهروي. (٢٠٠١). *تهذيب اللغة*. تحقيق: محمد عوض مرعب. دار إحياء التراث، بيروت. ط ١.
- ٦ الثعالبي، أبو منصور. (٢٠٠٢)، *فقه اللغة وسر العربية*، تحقيق: عبد الرازق مهدي، إحياء التراث، ط ١
- ٧ حسان، تمام. (٢٠٠٦). *اللغة العربية معناها ومبناها*. عالم الكتب، طه
- ٨ حمزة، حسن. (٢٠٢١). *التعريف وقوائم الشبوع في المعجم اللغوي العام*. مجلة الممارسات اللغوية، الجزائر، مجلد ١٢، عدد ١، ص ص ٩-٤٢.
- ٩ خاطر، محمود رشدي (١٩٥٥). *قائمة المفردات الشائعة في اللغة العربية*. المركز الدولي للتربية الأساسية في العالم العربي، مصر.
- ١٠ دوزي، رينهات بيرت آن. (١٩٩٧). *تكملة المعاجم العربية*. نقله إلى العربية وعلق عليه، محمد سليم النعيمي، وزارة الثقافة والإعلام، الجمهورية العراقية، ط ١، ج ٨.
- ١١ دياب، محمد حافظ. (١٩٨٥). *جماليات اللون في القصيدة العربية*، مجلة فصول، الهيئة العامة للكتاب، القاهرة، مجلد ٥، عدد ٢.
- ١٢ الزبيدي، محمد بن محمد بن عبد الرزاق الحسيني. (١٩٧٩). *تاج العروس من جواهر القاموس*، تحقيق: عبد الكريم العزباوي، راجعه، عبد الستار أحمد فراج، التراث العربي، الكويت، ج ١٨.

- ١٣ طعيمة، رشدي أحمد. (١٩٨٥). دليل عمل في إعداد المواد التعليمية لبرنامج تعليم العربية، معهد اللغة العربية، جامعة أم القرى، مكة المكرمة، ص ١-٥٠٨.
- ١٤ عاقل، فاخر (١٩٥٣). المفردات الأساسية للقراءة الابتدائية. دمشق.
- ١٥ عبد الصبور، محمد أحمد. (٢٠١٦). الكفاءة التعريفية في معاجم المصطلحات اللسانية، جسور، القاهرة، الإصدار الرابع، يناير ٢٠١٦.
- ١٦ عبد الصبور، محمد أحمد. (٢٠١٨). الدلالات الشعورية لألفاظ الألوان: دراسة على ضوء اللسانيات النفسية الكمية اللون الأصفر في مدونة الشعر العربي الحديث نموذجًا. مجلة الدراسات الإنسانية والأدبية، كلية الآداب، كفر الشيخ. عدد ١٥، مجلد ٤، ص ٢٠٠٥-٢٠٦٩.
- ١٧ عبد المقصود، محمود عبد الله. (٢٠٠٠م). مفاريد الألفاظ في القرآن الكريم دراسة لغوية. ماجستير. كلية اللغة العربية، جامعة الأزهر، ص ١-٣٧٦.
- ١٨ عبده، داود عطية (١٩٧٩). المفردات الشائعة في اللغة العربية. المملكة العربية السعودية، الرياض: مطبوعات جامعة الرياض.
- ١٩ علم، يحيى مير. (٢٠٢١). إحصاء الأفعال العربية ورصدها في المعجم الحاسوبي تأصيل وتوجيه، مجلة اللغويات العربية التراثية، العراق، مجلد ١٩، ص ٢٧-٧٩. December 2021, Journal of Arabic Linguistics Tradition, pp. 27-79, Volume 19
- ٢٠ علي، إبراهيم محمد. (٢٠٠١). اللون في الشعر العربي قبل الإسلام، لبنان، ط ١.
- ٢١ عمر، أحمد مختار. (١٩٧٧). اللغة واللون، عالم الكتب، القاهرة، ط ٢.
- ٢٢ عمر، أحمد مختار. (٢٠٠٨). معجم اللغة العربية المعاصرة، عالم الكتب، القاهرة، ط ١.
- ٢٣ الفراهيدي، أبو عبد الرحمن الخليل بن أحمد. (١٩٨٠). كتاب العين، تحقيق: مهدي المخزومي، وإبراهيم السامرائي. دار ومكتبة الهلال، بيروت، ط ١.
- ٢٤ الفيفي، عبد الله، والرحيلي، سمير، والقحطاني، محمد، والعصيمي، عبد الرحمن، والغامدي، أيمن. (٢٠١٦). معجم المفردات الشائعة لمتعلمي العربية "شائع". المؤتمر الثاني حول الاتجاهات الحديثة في تدريس اللغة العربية المستوى الثاني في: الرياض، فبراير ٢٠١٦، ص ١-١٣.

https://www.researchgate.net/publication/309195345_mjm

[almfrdat alshayt lmtlmy alrbyt shay](#)

٢٥ الفيومي، أحمد بن محمد بن علي. (١٩٨٧). *المصباح المنير في غريب الشرح الكبير*، مكتبة لبنان، بيروت، ط١.

٢٦ الكجراتي، جمال الدين، محمد طاهر بن علي الصديقي الهندي الفتّي. (١٩٦٧). *مجمع بحار الأنوار في غرائب التنزيل ولطائف الأخبار*. مطبعة مجلس دائرة المعارف العثمانية، ط٣.

٢٧ اللجنة الدائمة للرصيد اللغوي (١٩٧٥). *الرصيد اللغوي الوظيفي للمرحلة الأولى من التعليم الابتدائي*. تونس: الهيئة الاستشارية للمغرب العربي في التربية والتعليم.

٢٨ مجمع اللغة العربية. (٢٠٠٤). *المعجم الوسيط*. إعداد: مصطفى، إبراهيم، والزيات، أحمد، وعبد القادر، حامد، والنجار، محمد. مكتبة الشروق الدولية، القاهرة، ط٤.

٢٩ وحدة البحوث والمناهج (١٩٨٣). *قائمة مكة للمفردات الشائعة*. إعداد لجنة من المؤلفين. المملكة العربية السعودية، مكة المكرمة: جامعة أم القرى.

ثانياً- المراجع الأجنبية.

30 Alam, Yahya Mir. (2021). Statistics of Arabic Verbs in the Computational Arabic Lexicon: Survey of Findings and Implications. *Journal of Arabic Linguistics Tradition*, December 2021, Volume 19, pp. 27-79)

31 Amsler, R. A. (1982). Computational lexicology: A research program. *Proceedings of the June 7-10, 1982, National Computer Conference*, 657-663.

<https://doi.org/10.1145/1500774.1500857>

32 Assunção, C & Araújo, C. (2019). *Entries on the history of corpus linguistics*. *Linha D'Água (Online)*, São Paulo, v. 32, n. 1, p. 39-57, Jan-Abril 2019.

<https://www.researchgate.net/publication/332725200>. DOI:

10.11606/issn.2236-4242.v32i1p39-57

33 Barkovska, O., & Ivanenko, Y. (2023). Ways to determine the range of keywords in a frequency dictionary for text classification. *Computer Systems and Information*

34 Başaran, Broa. (2022). Measuring Word Frequency in Language Teaching Textbooks using LexiTürk. *International Online Journal of Education and Teaching (IOJET)*, 9(1). 571-583.

35 Berlin, B., Kay, P. (1969). *Basic Color Terms*. University of California Press, Berkeley

36 Brysbaert, M., Buchmeier, M., Conrad, M, Jacobs, A, M., Bólte, J, And Böhl, A. 2011. The Word Frequency Effect: A Review of Recent Developments and Implications for the Choice of Frequency Estimates in German. *Experimental Psychology*. hogrefe econtent. VOL. 58, NO.

5. <https://econtent.hogrefe.com/doi/epdf/10.1027/1618-3169/a000123>

37 Buckwalter, T., & Parkinson, D. (2014). *A Frequency Dictionary of Arabic* (0 ed.). Routledge.

<https://doi.org/10.4324/9780203883280>

38 Byrd, R. J., Calzolari, N., Chodorow, M. S., Klavans, J. L., Neff, M. S., & Rizk, O. A. (1987). Tools and methods for computational lexicology. *Computational Linguistics*, 13(3-4), 219-240.

39 C.L. Hardin: Ch.11 of: Backhaus, Kliegl and Werner: color Vision: Perspectives from Different Disciplines. Walter de

Gruyter, Berlin and New York, 1998.

On:<http://aardvark.ucsd.edu/color/hardin2.html>

40 COATS, S. (n.d.). Double modals in contemporary British and Irish speech. *English Language & Linguistics*, 1–26.

41 Crystal, D., & Crystal, D. (2008). *A dictionary of linguistics and phonetics* (6th ed). Blackwell Pub.

42 Davies, M., & Gardner, D. (2013). *A Frequency Dictionary of Contemporary American English* (0 ed.). Routledge. <https://doi.org/10.4324/9780203880883>

43 Dedrick, Don. (2013). A Review of: Carole P. Biggam: *The semantics of colour: A historical approach*. Cambridge: Cambridge University Press. in: *Language*, volume 89, number 4. Pp. 950–977

44 Derocher, James E. & Miron, Murray S. & Patten, Sam M. & Pratt, Charles C. (1973). *The counting of words. A review of the history, techniques and theory of word counts with annotated bibliography*. Syracuse, NY: Syracuse University Research Corporation (Microfiche)

45 Ech-Charfi, A. (2023). Word frequency and lexical coverage in English and Arabic. *Journal of Applied Language and Culture Studies*, 6(3), pp. 1–19.

46 Evans, R., M. (1948). *An introduction to color*, New York: John Wiley; London: Chapman & Hall

47 Fillmore, C. J. (2011). *Corpus Linguistics or Computer-Aided Armchair Linguistics*. In Svartvik, J. (Ed.).

Directions in corpus linguistics. Berlin, New York, de Gruyter Mouton, pp. 35–60, p3.

48 Francis, W.N.; Kučera, H (1982). Frequency Analysis of English Usage: Lexicon and Usage. Boston: Houghton Mifflin.

49 Guthrie, Louise., Pustejovsky, James., Wilks, Yorick., Slaton, Brian M. (1996). Lexicons in Natural Language Processing, Communications of the ACM January 1996/Vol. 39, No. 1, pp 62–72

50 Halliday, M. A. K. (2008). Linguistic function and literary style: An inquiry into the language of William Golding's *The Inheritors*. In *The Language and Literature Reader*. Routledge.

51 Hardin, C. L. (2013). Berlin and Kay Theory. In R. Luo (Ed.), *Encyclopedia of Color Science and Technology* (pp. 1–4). Springer New York. https://doi.org/10.1007/978-3-642-27851-8_62-2

52 Ingria, R. J. P. (1986). Structuring the Lexicon. Tutorial Session Paper presented at the 24th Annual Meeting of the Association for Computational Linguistics. Columbia University, New York. New York

53 Kopotev, M. (2021). Some Thoughts on Corpus and General Linguistics. *Philological Class*, 26, 90–102. <https://doi.org/10.51762/1FK-2021-26-02-07>

54 Kopotev, M. (2023). *Corpus Linguistics in Finland: a resource survey*.

55 Kuta, M., Chrzaszcz, P., & Kitowski, J. (2009). Increasing quality of the Corpus of Frequency Dictionary of

- Contemporary Polish for morphosyntactic tagging of the Polish language. *Computing and Informatics*, 28(3), 319–338.
- 56 Landau, J. M. (1959). A word count of modern Arabic prose, New York, American council of learned societies.
- 57 Levinson, S. C. (2001). *Yeli Dnye and the Theory of Basic Color Terms*. *Journal of Linguistic Anthropology* 10(1). pp.3–55. Copyright © 2001, American Anthropological Association.
- 58 Lyashevskaya, O. (2013). Frequency Dictionary of Inflectional Paradigms: Core Russian Vocabulary. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2286096>
- 59 Marcin Kuta, Paweł Chrzyszcz, Jacek Kitowski. *M. Kuta, P. Chrzyszcz, J. Kitowski*, Increasing quality of the corpus of frequency dictionary of contemporary Polish for morphosyntactic tagging of the Polish language. *Computing and Informatics*, Vol. 28, 2009, 319–338, p319.
- 60 Mark Davies. (2005) Selected Proceedings of the 7th Hispanic Linguistics Symposium, ed. David Eddington, PP. 106–115. Somerville, MA: Cascadilla Proceedings Project.
- 61 McEnery, T., Younis, N., & Hardie, A. (Eds.). (2018). *Arabic Corpus Linguistics*. Edinburgh University Press. <https://www.cambridge.org/core/books/arabic-corpus-linguistics/9740E03514B525289B7782CF0EDC791B>
- 62 Meyer, I., Skuce, D., Bowker, L., & Eck, K. (1992). Towards a New Generation of Terminological Resources: An Experiment in Building a Terminological Knowledge Base. *COLING 1992 Volume 3: The 14th International Conference on*

Computational Linguistics. COLING 1992.

<https://aclanthology.org/C92-3146>

63 Miron, Murray S. & Pratt, Charles C. (1973). *Manual for the development of language frequency counts*. Syracuse, NY: Syracuse University Research Corporation.

64 *Nation, I.S.P. (1990) Teaching and Learning Vocabulary*. Newbury House, New York. – *References—Scientific Research Publishing*. (n.d.). Retrieved June 15, 2023, from

[https://www.scirp.org/\(S\(lz5mqp453edsnp55rrgjt55\)\)/reference/ReferencesPapers.aspx?ReferenceID=1590732](https://www.scirp.org/(S(lz5mqp453edsnp55rrgjt55))/reference/ReferencesPapers.aspx?ReferenceID=1590732)

65 Rosch, Eleanor: *The Nature of Mental Codes for Color Categories*. In: *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 1(4), Nov 1975, 303–322.

66 Schmidt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave: Macmillan, pp. 3–43. DOI: [10.1057/9780230293977](https://doi.org/10.1057/9780230293977)

67 Spinelli, Tommaso. (2022). *The Diachronic Frequency of Latin Words*. A Computational Dictionary. Volume I (A/AB–AZYMVS). Propylaeum, 2022. pp. xiii– xlvi
.DOI: <https://doi.org/10.11588/propylaeum.1000>

68 Thomas Berg (2014): *On the Relationship between Type and Token Frequency*, *Journal of Quantitative Linguistics*, Vol. 21, No. 3, 199–222,

<http://dx.doi.org/10.1080/09296174.2014.911505>, <https://www.researchgate.net/publication/263287899>

69 Wang, J., Yuan, J., & Liu, W. (2023). *Sentiment Analysis and Corpus: Cognitive Perspective and Overhead–*

Accuracy Tradeoff. *ACM Transactions on Asian and Low-Resource Language Information Processing.*

70 Wanscher, j. H., kornerup, A. (1978). Methuen handbook of colour. Published by Eyre Methuen, Holt Art Books, Birmingham, United Kingdom

71 Xu, H. (2012). A critique of the controlled defining vocabulary in Longman Dictionary of Contemporary English. *Lexikos*, 22(1), 367–381.

ثالثاً - مواقع شبكة الانترنت.

72 <http://www.languagesgulper.com/eng/Papuan.html>

73 <https://poetry.dctabudhabi.ae/#/>

74 <https://shamela.ws/>

75 <https://tashkeel.alsharekh.org>

76 <https://www.bibalex.org/ica/ar/about.aspx>