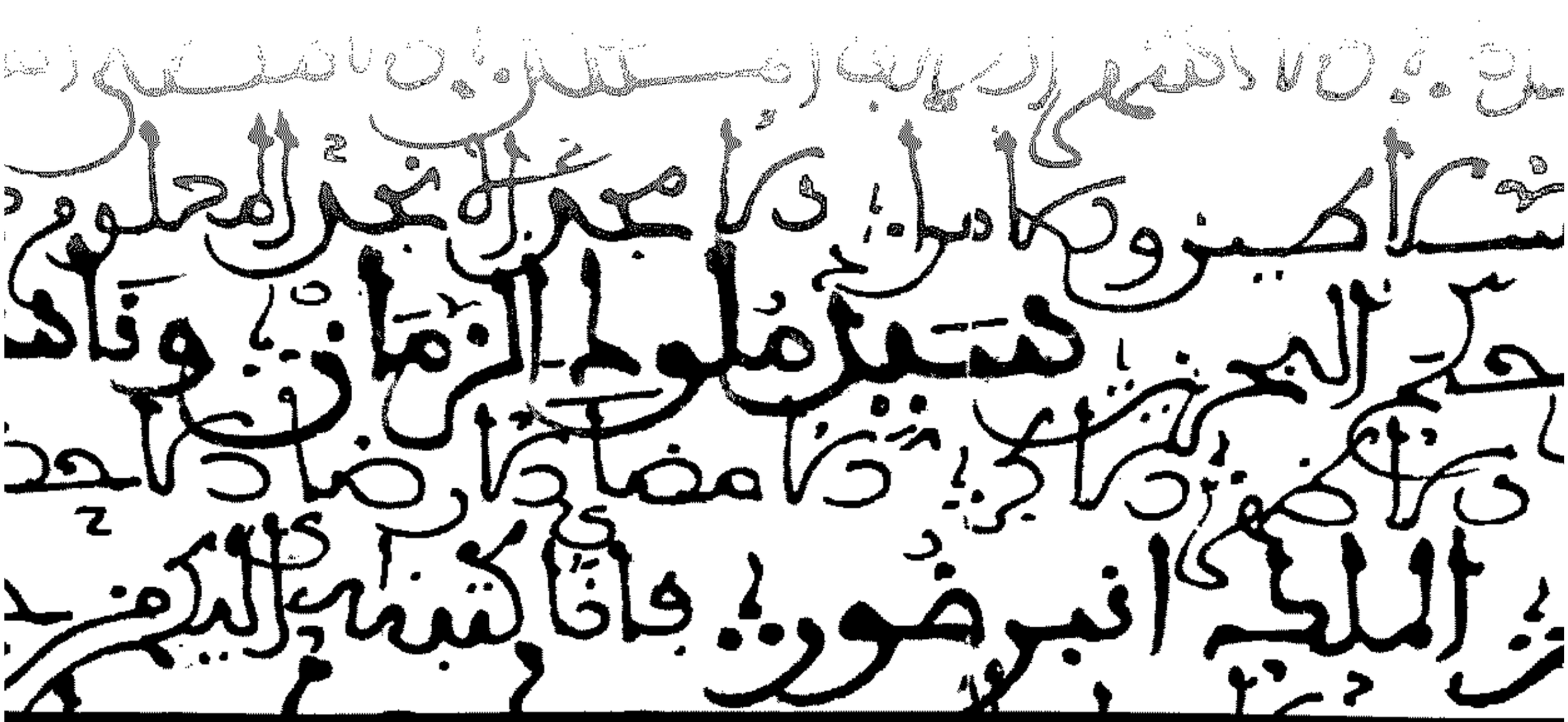
سلسلة المعرفة اللسانية
Lingustic Knoledge Series
LKS

الجزء الثاني

المعجمية العربية

قضايا وآفاق



إعداد وتقديم
د. منتصر أمين عبد الرحيم
د. حافظ إسماعيلي علوي

كنوز المعرفة
www.darkonoz.com

سلسلة المعرفة اللسانية
Linguistic Knowledge

يتأسس إنتاج المعرفة في الخطاب اللساني المعاصر على مبدأ تخريط المعرفة؛ أي مبدأ التداخل والتكامل بين اللسانيات وأنساق معرفية لها استقلاليتها الأنطولوجية في خريطة العلوم الحديثة.

وتأتي هذه السلسلة لتتفتح على أعمال تقرن الخطاب اللساني بعلوم متنوعة وبمحاور تطبيقية مختلفة مستجدة، لذلك سيتم التركيز على بعض القضايا التي لم يحصل فيها تراكم في سوق الكتابة اللسانية العربية.

ترحب السلسلة بنشر إسهامات الباحثين، سواء كانت دراسات وبحوث جماعية، أو كتب فردية.

من محاورنا القادمة:

- ❖ التخطيط اللساني والعولمة
- ❖ المعرفة اللسانية والأمراض اللغوية
- ❖ الخطاب اللساني المعاصر ووجائهه
- ❖ آفاق المعرفة اللسانية المعاصرة
- ❖ اللسانيات والعلوم المعرفية
- ❖ اللسانيات التطبيقية
- ❖ اللسانيات التربوية

المشرف العام:

الدكتور عبد القادر الفاسي الفهري

التحرير والتنسيق:

د. حافظ إسماعيلي　　د. امحمد الملاخ

د. منتصر أمين　　د. امحمد إسماعيلي

العنوان الإلكتروني:

knowledgelinguistic@gmail.com

المعجميـــة العربيـــة
قضايـــا وآفـــاق

مجموعة من المؤلفين

إعداد وتقديم

د. منتصر أمين عبد الرحيم د. حافظ إسماعيلي علوي

الجزء الثاني

الطبعة الأولى

1435هـ - 2014م

المملكة الأردنية الهاشمية
رقم الإيداع لدى دائرة المكتبة الوطنية: (2013/11/4087)

413,28
القهري، عبدالقادر الفاسي
المعجمية العربية/ قضايا وآفاق / عبدالقادر الفاسي الفهري، حافظ إسماعيلي علوي. – عمان: دار كنوز المعرفة للنشر والتوزيع، 2013
(346) ص.
ر.إ.: 2013/11/4087.
الواصفات: /اللغة العربية // القواميس /

يتحمل المؤلف كامل المسؤولية القانونية عن محتوى مصنفه ولا يعبّر هذا المصنف عن رأي دائرة المكتبة الوطنية أو أي جهة حكومية أخرى.

ردمك : ISBN: 978 - 9957 - 74 - 321 - 5

حقوق النشر محفوظة

جميع الحقوق الملكية والفكرية محفوظة لدار كنوز المعرفة- عمان- الأردن، ويحظر طبع أو تصوير أو ترجمة أو إعادة تنفيذ الكتاب كاملا أو مجزءا أو تسجيله على أشرطة كاسيت أو إدخاله على كمبيوتر أو برمجته على اسطوانات ضوئية إلا بموافقة الناشر خطيا

دار كنوز المعرفة العلمية للنشر والتوزيع

الأردن - عمان - وسط البلد - مجمع الفحيص التجاري
تلفون: 4655877 6 962+ - فاكس: 4655875 6 962+
موبايل: 5525494 79 962+ -ص. ب 712577 عمان
الموقع الإلكتروني: www.darkonoz.com
إيميل: info@darkonoz.com - dar_konoz@yahoo.com

المشاركون في الكتاب

د. أشرف عبده	مصر
د. امحمد الملاخ	المغرب
د. بولا سانتيان غريم	إسبانيا
د. جورج متري عبد المسيح	لبنان
د. حافظ إسماعيلي علوي	المغرب
د. خالد اليعبودي	المغرب
د. ربيعة العربي	المغرب
د. عبد الرحمن بودرع	المغرب
د. عبد الرزاق بنور	تونس
د. عبد العلي الودغيري	المغرب
د. عبد الفتاح الفرجاوي	تونس
د. عبد القادر الفاسي الفهري	المغرب
د. عبد القادر سلامي	الجزائر
د. عز الدين البوشيخي	المغرب
د. علي القاسمي	العراق
د. فاتن الخولي	مصر
د. محمد خالد الفجر	سوريا
د. محمد خطابي	المغرب
د. محمد غاليم	المغرب
د. مختار درقاوي	الجزائر
د. مصطفى غلفان	المغرب
د. المعتز بالله السعيد	مصر
د. منتصر أمين عبد الرحيم	مصر
د. ميشال زكريا	لبنان
د. وفاء كامل فايد	مصر
د. وليد العناتي	الأردن
د. يوسف محمد أبوعامر	مصر

الفهرس

❖ التقديم		٩
المحور الثالث: نحو آفاق جديدة للمعجمية العربية		٢١
❖ هل يعد معجم الاستشهادات معجماً؟	د علي القاسمي	٢٣
❖ بناء المعجم التاريخي للغة العربية واقتضاءاته النظرية	د عز الدين البوشيخي	٣٧
❖ مادة المعجم التاريخي للغة العربية	د عبد الرحمن بودرع	٤٩
❖ بنية المعجم العربي واستخدامه بين البشر والآلة	د يوسف محمد أبو عامر	٨٧
❖ المعجم التكراري لألفاظ القرآن الكريم: المنهج والنموذج	د المعتز بالله السعيد	١٤٥
❖ اللغة بين المعجم والحاسوب: بناء المعجم الإسلامي من منظور لساني حاسوبي	د فاتن الخولي	١٧٧
المحور الرابع: من قضايا المعجم		١٩٩
❖ إنشاء قاعدة معجمية عربية مولدة	د عبد القادر الفاسي الفهري	٢٠١
❖ المعجم الذهني والتقييس الحاسوبي	د امحمد الملاخ ود حافظ إسماعيلي علوي	٢٣٥
❖ نحو تصور جديد لتخصيص المعاني المعجمية	د محمد غاليم	٢٥٥
❖ السمات الزمنية والجهية والوجهية وسمها ومعجمتها	د امحمد الملاخ	٢٧٥
❖ تصنيف مجدَّد ومجدِّد للمتلازمات اللفظية في العربية	باولا سانتيان غريم	٢٩٧
❖ في أسبقية المصطلح على الكلمة	د عبد الرزاق بنور	٣٢٧

بنية المعجم العربي
واستخدامه بين البشر والآلة

يوسف أبو عامر

أحدث اختراع الحاسب الآلي نقلة نوعية في الأبحاث والدراسات العلمية في مختلف المجالات والتخصصات. ولم تكن الدراسات اللغوية -ومنها المعجمية- بمنأى عن الإفادة من قدراته الهائلة؛ إذ عُدَّت الإفادة من الحاسب الآلي في صناعة المعجم من أهم مجالات علم اللغة الحاسوبي *Computational Linguistics*، وأكثرها تلبية للمتطلبات العلمية والثقافية في العالم المعاصر. لقد ارتبط هذا الاتجاه نحو حوسبة المعجم -ارتباطًا واضحًا- بمجموعة المميزات التي تقدمها التقنية لكل الأطراف ذات الصلة بالمعاجم مقارنة بنظائرها الورقية، فكان بناء القواميس المقروءة آليًا *Machine Readable Dictionaries*. ومع ذلك ظلت عملية بناء المعاجم -حتى في صورتها الآلية- موجهة بالأساس إلى المستخدم البشري، وبالتالي ظلت هناك مشكلات جوهرية تعانيها المعاجم الحاسوبية العربية، لا سيما عند استخدامها في تطبيقات المعالجة الآلية للغة العربية *Arabic Natural Language Processing*.

إن بناء المعاجم الموجهة للمستخدم البشري يختلف -اختلافًا واضحًا- عن بناء المعاجم التي تستخدم في تطبيقات المعالجة الآلية؛ لأنه رغم اشتمال الأولى على قدر من المعلومات القيمة، فإنها لا تصلح للأخيرة؛ إذ تحتاج الأخيرة إلى معلومات تفصيلية على المستويات الصرفية، والنحوية، والدلالية؛ لأن المستخدم في هذه الحالة هو الآلة، وليس الإنسان. ولهذا تحتاج المعاجم أن تكون أكثر شموليةً، وتنظيمًا، ووضوحًا، وتماسكًا، وشكليةً من معاجم الاستخدام البشري.

تهدف الدراسة -على مستوى التأسيس النظري- إلى تأصيل منهجي لبناء المعجم الحاسوبي العربي؛ على نحوٍ يجعله وافيا بمتطلبات المعالجة الآلية للغة

العربية المعاصرة؛ إذ تقدم تصورًا للأساس الذي يمكن أن يبنى عليه المعجم من الزاويتين: اللغوية والحاسوبية، كما تقدم منهجية لتمثيل المحتوى اللغوي للمداخل المعجمية. أما على مستوى التطبيق فتقدم تصميمًا لمعجم آلي مبني على أساس الجذع اللغوي، ومزوَّد بمحلل صرفي آلي ثنائي الاتجاه، ويحتوي على بيانات تفصيلية للمعلومات اللغوية للمداخل المعجمية.

تنقسم هذه الدراسة إلى ستة أجزاء: يتناول الأول منها -بإيجاز- رحلة المعجم العربي من الورقية إلى الآلية، في حين يعرض الثاني أبرز الاختلافات بين بناء المعاجم التقليدية الموجهة للمستخدم البشري ومعاجم نظم المعالجة الآلية، ويقدم الثالث إجابة عن السؤال المتعلق بالأساس الذي يجب بناء المعجم الحاسوبي العربي عليه؛ أي على أساس الجذور والأوزان، أم على أساس الجذوع اللغوية، كما يتناول مجموعة الخصائص الصرفية النحوية والدلالية التي تُرَمَّز في المعجم، ويتضمن الجزء الرابع توصيفًا للمدونة التي اعتمد عليها الباحث في استقاء المداخل المعجمية، وأخيرًا يقدم الجزء الخامس توصيفًا للمعجم الآلي الذي صممه الباحث بغرض استخدامه في تطبيقات المعالجة الآلية للعربية المعاصرة، وأخيرًا تعرض الخاتمة النتائج التي توصلت إليها الدراسة.

رحلة المعجم العربي من الورقية إلى الآلية:

ارتبطت نشأة المعاجم -بوجه عام- عند كثير من الأمم ارتباطًا وثيقًا بالعامل الديني؛ فمنذ الألفية الثانية قبل الميلاد كان للدوافع الدينية الأثر الفاعل في نشأة المعاجم وتأليفها. لقد كان الغرض الأول من تأليف المعاجم عند الهنود هو فهم اللغة السنسكريتية[1] *Sanskrit*؛ لأنها كانت لغة الأغاني والنصوص المقدسة،

---

(١) اللغة السنسكريتية هي لغة التراث والثقافة والنصوص المقدسة عند الهنود، وقد اكتسبت أهميةً كبيرةً، واهتمامًا واسعًا؛ نظرًا لظروف اكتشافها، وارتباط ذلك بنشأة اللغويات التاريخية *Historical Linguistics*. انظر:

- *Brockington, J. (2006). Sanskrit. In Brown, K. & Ogilvie, S. (ed). (2009). Concise Encyclopedia of Languages of the World. First edition. Elsevier Ltd; Oxford, UK. P. 918.*

كما ارتبطت حاجة الصينيين إلى المعاجم برغبتهم في تعرُّف أعمال كونفوشيوس.[1] وفي السياق ذاته، كانت حاجة الأوروبيين لتعليم القساوسة لغة الإنجيل كامنة وراء حاجتهم إلى المعاجم.[2]

ولم تكن نشأة المعاجم عند العرب استثناءً؛ إذ كانت حراسة القرآن من الخطأ في النطق أو الفهم أهم الأسباب التي مهدت لنشأة المعجم العربي، كما اقترن هذا السبب ببعض الأسباب الأخرى مثل زحف حياة الحواضر على البدو، مما هدد بنضوب المعين الذي كان يستقي منه الرواة، وكذلك حاجة الرواة إلى تسجيل رواياتهم اللغوية، بالإضافة إلى خوفهم على اللغة من الانقراض بانقراض الحافظين لها.[3]

لقد مَرَّ المعجم العربي برحلة طويلة بدأت إرهاصاتها منذ القرن الأول للهجرة، حين حاول الصحابة -ومن بعدهم اللغويون- تفسير غريب القرآن ومشكله. وفي هذا الصدد نجد كتابًا منسوبًا لابن عباس بعنوان «في غريب

---

(١) فيلسوف صيني ولد عام ٥٥١ قبل الميلاد، في ضاحية صغيرة تسمى «لو *Lu*»، لأسرة أرسطوقراطية. ويعد كونفوشيوس أحد رواد الفكر الإنساني، وصاحب مدرسة كونفوشيوس التي تركت أثرًا عميقًا في التاريخ البشري، انظر:

- رسلان، صلاح بسيوني. (د. ت.). كونفوشيوس رائد الفكر الإنساني. نسخة إلكترونية. ص ٦٠ وما بعدها.

- *Yao, X. (2000). An Introduction to Confucianism. Cambridge: Cambridge University Press. P.16.*

(2) Sterkenburg, P. The Dictionary: Definition and History. In Sterkenburg, P. (2003). A Practical Guide to Lexicography . Terminology and Lexicography Research and Practice. ISSN 7067; V. 6. John Benjamin B.V. P. 9.

(٣) أبو سكين، عبدالحميد محمد. (١٩٨١). المعاجم العربية: مدارسها ومناهجها. الطبعة الثانية. القاهرة: الفاروق الحديثة للطباعة والنشر. ص١٧ - ١٨.

القرآن»،[1] ارتأى حسين نصار أن «هذا الكتاب كان يضم بعض الأقوال التي أدلى بها ابن عباس في تفسير الغريب من ألفاظ القرآن، وأنه لم يكن هو الذي دونها، وإنما بعض رواة هذه الأقوال».[2]

أما الرحلة الفعلية للمعجم العربي، فقد انطلقت من قاعدة «الرسائل الصغيرة» التي اختص كل منها بمفردات موضوع معين، ولم تُرتب مفرداتها ترتيبًا ألفبائيًا، وإنما رُتِّبَت تبعًا للمعنى، أو شكل الكلمة، أو ترتيبًا عشوائيًا. وقد خصص حسين نصار كتابًا كاملًا من مؤلفه «المعجم العربي: نشأته وتطوره» لهذه الرسائل، سمَّاه «الرسائل اللغوية على الموضوعات»، وقسمه إلى تسعة أبواب: الغريبين والفقه، واللغات والعامي والمعرَّب، والهمز، والحيوان، والنوادر، والبلدان والمواضع، والإفراد والتثنية والجمع، والأبنية، والصفات.[3]

لكن جمع المفردات على هذا النحو؛ أي باعتبار المعنى، لم يكن المنهج الأوحد الذي اتبعه العرب؛ إذ إن ثمة منهجًا آخر عرفته العرب يقوم على جمع المفردات باعتبار اللفظ. وقد كان للخليل بن أحمد الفراهيدي فضل السبق في هذا المنحى؛ فقد ابتكر أسلوبًا يؤدي إلى جمع مفردات العربية وتدوينها بين دفتي كتاب، ووضع نهجًا لهذا الجمع يقوم على أسس رياضية بحتة، ومن ثم إلى بناء معجم يضم ألفاظ العربية.[4]

وكان منهج الخليل يهدف إلى استقصاء كلمات العربية اعتمادًا على أساسين: الأول انحصار الحروف في تسعة وعشرين حرفًا، وانحصار الكلمات

---

(١) يشير أحمد الشرقاوي إقبال في «معجم المعاجم» إلى وجود مخطوطة الكتاب بالظاهرية، وبمكتبة طلعت بدار الكتب المصرية، وبرلين، انظر: إقبال، أحمد الشرقاوي. (١٩٩٣). معجم المعاجم. الطبعة الثانية. بيروت: دار الغرب الإسلامي. ص٧.

(٢) نصار، حسين. (١٩٨٨). المعجم العربي: نشأته وتطوره. الطبعة الرابعة. الجزء الأول. القاهرة: دار مصر للطباعة. ص ٣٣.

(٣) السابق نفسه. ص ٣٣ – ١٧٠.

(٤) الخطيب، عدنان. (١٩٩٤). المعجم العربي بين الماضي والحاضر. الطبعة الثانية. لبنان: مكتبة لبنان ناشرون. ص ٣٥.

العربية بين الثنائي والخماسي، ومن ثم يمكن حصر اللغة من خلال تتبع دوران كل حرف في كل بناء من هذه الأبنية.[١] وعلى هذين الأساسين قام الخليل بجمع مفردات معجمه الشهير بـ «كتاب العين»؛[٢] ليكون ذلك إيذانًا بدخول العرب مرحلة التأليف المعجمي الحقيقي؛ إذ يعد العين «أول معجم لغوي مرتب موسع عرفته العرب».[٣]

ثم نشطت حركة التأليف المعجمي بعد الخليل مباشرة، وتعددت مدارس ترتيب المعاجم؛ فالتزم البعض منهج الخليل في ترتيب الحروف بحسب المخارج، كالقالي في «البارع»، والأزهري في «التهذيب»، والصاحب بن عباد في «المحيط»، وأهمل البعض الآخر هذا المنهج، والتزم الترتيب الألفبائي وفقًا للحروف الأصلية الأولى من الكلمة، كابن دريد في «الجمهرة»، وابن فارس في «المقاييس» و«المجمل»، والزمخشري في «أساس البلاغة»، أو اعتمادًا على الحرف الأخير، ثم الحرف الأصلي الأول فالثاني كالجوهري في «الصِّحاح»، وابن منظور في «اللسان»، والفيروزأبادي في «القاموس المحيط»، والزبيدي في «تاج العروس». وهناك فريق ثالث اعتمد الترتيب الموضوعي، أي باعتبار المعنى، دون التزام بترتيب الكلمات داخل المعنى الواحد، ومن أشهر أصحاب هذا الاتجاه ابن سيده في «المخصص»، والثعالبي في «فقه اللغة»، والجواليقي في «المُعرَّب». وقد مثّلت هذه المعاجم العربية الوسطية مرحلة هامة وفارقة في تاريخ المعاجم في العالم بأسره.

وفي العصر الحديث، ظهرت بعض المحاولات لتأليف معجم حديث يحافظ على سلامة اللغة، ويجعلها وافية بمطالب العلوم والفنون، وملائمة لحاجات

---

(١) نصار، حسين. (١٩٨٨). الجزء الأول. ص ١٧٥.

(٢) ثمة خلاف بين الدارسين حول نسبة العين للخليل بين رافضٍ، ومؤكدٍ، ومتحفظٍ، ومترددٍ، وشاكٍّ. ولسنا هنا بصدد الدخول إلى معترك هذا الخلاف، لكن للبحث حوله يمكن الرجوع إلى: نصار، حسين. (١٩٨٨). الجزء الأول. ص ٢١٩ – ٢٣٢.

(٣) الباتلي، أحمد بن عبد الله. (١٩٩٢). المعاجم اللغوية وطرق ترتيبها. الرياض: دار الراية. الطبعة الأولى. ص ١٥.

الحياة في العصر الحديث، وكان من أبرز نتاج هذه المحاولات صدور المعجم الوسيط عن مجمع اللغة العربية بالقاهرة، وقد صدرت طبعته الأولى سنة ١٩٦٠م.(١)

وقد أحدث اختراع الحاسب الآلي نقلة نوعية في الأبحاث والدراسات العلمية في مختلف المجالات والتخصصات. ولم تكن الدراسات اللغوية -ومنها المعجمية- بمنأى عن الإفادة من قدراته الهائلة؛ إذ عُدَّت الإفادة من الحاسب الآلي في صناعة المعجم من أهم مجالات علم اللغة الحاسوبي، وأكثرها تلبية للمتطلبات العلمية والثقافية في العالم المعاصر.(٢)

وقد ظهرت فكرة الاستعانة بالحاسب الآلي -بوجه عام- في صناعة المعجم منذ فترة مبكرة، وذلك عندما «اتضحت فكرة إمكان تحويل الوجود المادي للغة - بوصفها تتابعًا صوتيًّا منطوقًا- إلى نظام آخر من التتابعات على أساس البطاقات المثقبة -مثلًا- دون أن يؤدي ذلك إلى خسارة في نقل المعلومات بالنسبة للحاسب الآلي.»(٣) لقد انطلق العمل الحاسوبي في المعاجم مع بداية ستينات القرن الماضي؛ فقد أجريت بعض الأعمال الحاسوبية على القواميس ومستودعات المفردات *Thesauruses*، ومع بداية الثمانينات اكتسبت القواميس المقروءة آليًّا *Machine Readable Dictionaries* اهتمام الباحثين في تخصصات عديدة؛ في محاولة للإفادة من المعلومات اللغوية الموجودة في هذه القواميس لبناء القواميس الحاسوبية، وقواعد المعرفة المعجمية التي تحتاج إليها الحاسبات لمعالجة اللغة الطبيعية.(٤)

---

(١) مجمع اللغة العربية بالقاهرة. (٢٠٠٤). تصدير: المعجم الوسيط. الطبعة الرابعة. القاهرة: مكتبة الشروق الدولية.

(٢) حجازي، محمود فهمي. (د.ت). البحث اللغوي. القاهرة: دار غريب للطباعة والنشر. ص ٧١.

(٣) السابق نفسه. الصفحة نفسها.

(4) Kruyt, T. Multifunctional Linguistic Databases: Their Multiple Use. In Sterkenburg, P. (ed). (2003). P. 195.

وقد أسهم توفر المادة اللغوية في الحاسوب وإمكانية تحليها رقميًا وتحديد المهمل والمندثر واكتشاف الكلمات الجديدة في عمل مشروعات كبيرة على النحو الذي يظهر في مشروع كوبيلد *COBUILD* اختصارا *Collins Birmingham University International Language Database*، وهو مشروع مشترك بين جامعة بيرمنجهام ودار كولينز للنشر.

انطلق مشروع كوبيلد عام ١٩٨٠م وتحققت ثمرته عام ١٩٨٧م بإصدار الطبعة الأولى لقاموس كولينز كوبيلد للغة الإنجليزية *Collins COBUILD English Language Dictionary*، ثم توالت الطبعات الثانية سنة ١٩٩٥م، والثالثة سنة ٢٠٠١، والرابعة سنة ٢٠٠٣م، والخامسة سنة ٢٠٠٦م. وكان الإسهام الأبرز لهذا المشروع هو بناء مدونة للنصوص اللغوية المعاصرة (بنك اللغة الإنجليزية *Bank of English*) وتحليها، وبالتالي قدم المعجم مداخل تمثل الاستخدام الفعلي للغة.(١)

ومع زيادة الاهتمام بالمعالجة الآلية للغات الطبيعية، أصبح المعجم ركيزةً أساسيةً يعتمد عليها بناء النظم المختلفة للمعالجة الآلية للغات الطبيعية كالترجمة الآلية *Machine Translation*، واسترجاع المعلومات *Information Retrieval*، والتلخيص الآلي *Automatic Summarization*، والتحليل الصرفي *Morphological Analysis* ...إلخ.

أما المعجم العربي، فقد بدأ دخول عصر الحاسوب في سبعينيات القرن الماضي ليبدأ مرحلة جديدة بتحديات جديدة، وآمال عريضة في مواكبة التطور الكبير في الدراسات المعجمية في أوروبا وأمريكا. وقد بدأ المعجم العربي دخول عصر الحاسوب بدراسة إحصائية لجذور اللغة العربية المدونة بمعجم الصحاح، بالإضافة إلى الحروف الداخلة في تركيب تلك الجذور. وهي دراسة أجريت بجامعة الكويت، وقام بها علي حلمي موسى -أستاذ الفيزياء بجامعة الكويت آنذاك- بناء على اقتراح قدمه إبراهيم أنيس؛ إذ استخدم الحاسب الآلي

---

(1) Sinclair, J. (ed.) 1987. Looking Up: An Account of the COBUILD Project in Lexical Computing. London: Collins.

لإحصاء الجذور الثلاثية وغير الثلاثية لمعجم الصحاح.[1]

وانتهت هذه الدراسة إلى «عرض جذور مفردات اللغة العربية المستخدَمة والواردة في المعجم في صورة جداول توضح تردد الحروف في المواقع المختلفة من الكلمة، وتتابع الحروف، واختلافها من حرف لآخر».[2] وقد تبع ذلك دراسات أخرى لم تخرج عن هذا الإطار الإحصائي، منها دراسة إحصائية لجذور معجم لسان العرب لابن منظور عام ١٩٧٣م، ودراسة أخرى لإحصاء جذور معجم تاج العروس للزبيدي عام ١٩٧٣م.

لقد ارتبط هذا الاتجاه نحو حوسبة المعجم العربي -ارتباطا واضحًا- بمجموعة المميزات التي تقدمها التقنية لكل الأطراف ذات الصلة بالمعاجم مقارنة بنظائرها الورقية. تعد المرونة أولى مميزات المعاجم الإلكترونية التي تتعلق بمستخدم المعجم؛ إذ تتضح أهمية هذه الميزة من جهتين: الأولى تتعلق بآليات، أو طرق الوصول إلى المعلومات داخل القاموس. ففي الوقت الذي تسمح فيه القواميس الورقية بطريقة بحث أحادية -وفقًا للترتيب الأبجدي في أغلب الأحيان- يمتلك المستخدم طرقًا بحثية متعددة للوصول إلى المعلومات في حالة استخدام القاموس الإلكتروني. أما الجهة الثانية فترتبط بقابلية المعجم لإجراء التعديلات التي تتماشى مع التطور اللغوي؛ فالقواميس الورقية تتسم -في هذا الصدد- بنوع من الجمود يجعلها تتوقف عند لحظة النشر، في حين تتيح القواميس الإلكترونية إجراء التعديلات التي تتماشى مع التطورات اللغوية، على نحو يسمح بإضافة المفردات الجديدة، وتصحيح الأخطاء بشكل دائم ومتجدد.[3]

تتمثل الميزة الثانية من مميزات القواميس الإلكترونية في سرعة الوصول إلى المعلومات؛ إذ تتيح تقنيات البحث الآلي الوصول إلى المعلومات الخاصة

---

(١) موسى، علي حلمي. (١٩٧٨). مقدمة: دراسة إحصائية لجذور معجم الصحاح باستخدام الكمبيوتر. الهيئة المصرية العامة للكتاب.

(٢) السابق نفسه. ص ١٠.

(3) Sterkenburg, P. (2003). P. 5 -6 .

بالمدخل المعجمي بمجرد النقر على كلمة على الشاشة.[1] أما الإمكانات التي تتيحها التكنولوجيا للمعجمي، فقد رصد عمر مهديوي عددًا منها في دراسته عن المعجم العربي في ضوء اللسانيات الحاسوبية، ومن هذه الإمكانات[2]:

- تحرير النصوص باستخدام الآلة.
- استرجاع النصوص بسهولة وسرعة.
- انخفاض تكاليف التخزين.
- استخدام الخوارزميات في معالجة النصوص.
- الاستفادة من المدققات الإملائية والنحوية في مراجعة النصوص.
- دراسة تطور المخزون اللغوي عبر الزمن.

ويبدو أن المعجم العربي -اتكاءً على هذه المميزات- قد بدأ يأخذ -منذ الثمانينات- سمتًا حاسوبيًا من خلال تشييد قواعد البيانات المعجمية *Lexical Databases*، والمعاجم الآلية؛ كي تُستخدم في المعالجة الآلية للغة العربية، أو لاستخدامها من قبل المستخدم البشري. كما تزايد الاهتمام بمعالجة المعجم العربي آليًا داخل الوطن العربي وخارجه، سواء أكان هذا الاهتمام جامعيًا أكاديميًا، أو في شكل مؤسسات تجارية.[3]

---

(1) Ibid. P. 7.

(٢) مهديوي، عمر. (٢٠٠٧). المعجم العربي في ضوء اللسانيات الحاسوبية. الندوة الدولية الأولى عن الحاسب واللغة العربية. مدينة الملك عبدالعزيز للعلوم والتقنية. الرياض، الفترة من ١٠ - ١٢ نوفمبر ٢٠٠٧.

(٣) أشارت وفاء كامل في دراستها عن واقع المعجمية الحاسوبية العربية إلى بعض المؤسسات المعنية بالعمل المعجمي الحاسوبي العربي داخل الوطن العربي وخارجه، ومن هذه المؤسسات جامعة عمان، وجامعة محمد الخامس، وجامعة بالماند لبنان، والشركة الهندسية لتطوير أنظمة الحاسب (*RDI*)، وشركة صخر. ومن خارج الوطن العربي جامعة كوبنهاجن، وجامعة ليون ٢، وجامعة أترخت، وغيرها. راجع:

- فايد، وفاء كامل. (٢٠٠٧). واقع المعجمية الحاسوبية العربية. ندوة (العربية في عصر العولمة)، المجلس الأعلى للثقافة، ومجمع اللغة العربية بالقاهرة (في إطار الاحتفال باليوبيل الماسي لمجمع اللغة العربية)، القاهرة، مارس ٢٠٠٧.

وعلى الرغم من الاهتمام الذي اكتسبته المعالجة الآلية للمعجم العربي، فإن ثمة مشكلات جوهرية تعانيها المعجمية الحاسوبية العربية، تعالج الدراسة أهمها؛ إذ اعتمدت «معظم المصادر المعجمية المحوسبة على المعاجم التقليدية، التي تفتقر بدورها إلى الأسس النظرية البحثية التي تجعلها مادة كافية لمطالب معالجة اللغات الطبيعية»، وهو أمر راجع بدوره إلى عدم التمييز بين أسس بناء المعاجم المصممة للمستخدم البشري ونظيراتها المصممة بغرض استخدامها في نظم المعالجة الآلية.(١) وتتجاوز الدراسة هذه المشكلة من خلال بناء معجم آلي ثنائي اللغة، قادر على تلبية متطلبات نظم المعالجة الآلية للغة العربية بمستوياتها المختلفة، كما يعتمد المعجم في مصادره المعجمية على مادة المدونة اللغوية المحوسبة، لا على المعاجم التقليدية.

الاختلافات بين بناء المعاجم التقليدية ومعاجم نظم المعالجة الآلية:

تختلف المعاجم من حيث طبيعتها إلى معاجم ورقية، وأخرى آلية، ومن حيث الوظيفة المنوطة بها إلى معاجم مصممة للمستخدم البشري، وأخرى الهدف منها توصيل المعلومات إلى النظم التي تقوم بتحليل اللغة الطبيعية -النصوص أو الكلام- وإنتاجها. وقد ترتب على اختلاف الوظيفة -أو الهدف من البناء- سؤال بديهي يتعلق بمدى كفاية المعلومات في المعاجم المصممة للمستخدم البشري لتلبية متطلبات نظم المعالجة الآلية.

إن المعاجم الآلية المصممة للمستخدم البشري لا تختلف -في جوهرها- عن القاموس الورقي المطبوع، «بيد أن المعلومات تكون موجودة في ملفات نصيَّة *Text files*:»(٢) لأن المستهدف في الحالتين واحد، بل إن بعض هذه المعاجم الآلية لا تعدو -بالفعل- كونها نسخًا إلكترونية من قواميس ورقية، مع الأخذ في الاعتبار مجموعة المميزات التي يضيفها التصميم الآلي للمستخدم البشري.

والحقيقة أن هذا الأمر يختلف عن بناء معاجم تستخدم في تطبيقات

---

(١) السابق نفسه.

(2) Sterkenburg, P. (2003). P. 5.

المعالجة الآلية؛ لأنه على الرغم من اشتمال الأولى على قدر من المعلومات القيمة، فإنها لا تصلح للأخيرة؛ إذ تحتاج الأخيرة إلى معلومات تفصيلية على المستويات الصرفية، والنحوية، والدلالية، والسياقية؛ لأن المستخدم في هذه الحالة هو الآلة، وليس الإنسان. ولهذا تحتاج المعاجم أن تكون أكثر شمولية، وتنظيمًا، ووضوحًا، وتماسكًا، وشكليةً من معاجم الاستخدام البشري. وقد رُصِدَت أبرز الاختلافات بين بناء المعاجم التقليدية الموجهة للمستخدم البشري ومعاجم نظم المعالجة الآلية فيما يلي:(١)

وضوح التمثيلات المعجمية:

يعد وضوح التمثيل المعجمي *Lexical representation* من أهم مبادئ بناء المعجم الحاسوبية المصممة لتطبيقات المعالجة الآلية؛ فالمعرفة التي يحتوي عليها المعجم يجب أن تكون مُفَصَّلةً -بشكل كافٍ- بحيث تدعم نظم المعالجة، لا سيما النظم التي تعتمد على القواعد *Rule-based systems*، وهي نظم تعتمد على نظريات لغوية، مثل: النحو الوظيفي المعجمي *Lexical-functional grammar*، ونحو الاعتماد *Dependency syntax*، والنحو التصنيفي *Categorical grammar* ...إلخ.

شمولية البنية:

تعد المعاجم الحاسوبية في الحقيقة قواعد معرفة مركبة، ومن ثم يجب أن تكون البنى التي تربط الكلمات كلها لها أهمية البنى التي تربط عناصر الكلمات. كما يجب اعتبار التجمعات اللفظية -كالتصاحبات اللفظية *Collocations*، والتعابير الاصطلاحية *Idioms*، والأسماء المركبة *Compound Words* وغيرها- مداخل معجمية يتم تمثيلها داخل المعجم.

---

(1) Wilson, R. & Frank C. (eds.). (1999). Computational Lexicons. In The MIT Encyclopedia of the Cognitive Sciences. The MIT Press; Cambridge; Massachusetts; London; England. P. 161.

التغطية الكافية:

يجب أن يكون المعجم شاملًا مجال المعجم؛ فالمعجم الحقيقي يمكن أن يشتمل -مثلاً- على حوالي أربعمائة ألف (٤٠٠,٠٠٠) مدخل حسب التطبيق والهدف منه.

قابلية التقييم:

تُقَيَّم المعاجم الحاسوبية باعتبارات ثلاثة هي:

- التغطية *Coverage*: وتكون التغطية -أفقيًا- على مستوى المعجم، ورأسيًا على مستوى المعلومات المعجمية.
- القابلية للتوسيع *Extensibility*: ما مدى سهولة إضافة المعلومات إلى المدخل المعجمي؟ وإلى أي مدى يمكن أن تتسق المعلومات الجديدة مع البنى المعجمية الأخرى؟
- الاستخدام *Utility*: ما مدى فائدة المداخل المعجمية لتطبيقات المعالجة الآلية؟

تتطلب نظم المعالجة المعالجة الآلية للغات الطبيعية -من أكثرها بساطة إلى أكثرها تعقيدًا- معرفة بالمفردات؛ ففي معظم الأحوال تحتاج النظم -لكي تكون أكثر فاعلية وعملية- إلى تزويدها بمعجم أساسي يغطي الاستخدام الفعلي لمفردات اللغة، وكذلك المعلومات اللغوية المتعلقة بها. حتى مع التطبيقات البسيطة كالمدققات الإملائية قد لا يخلو الأمر من حاجة إلى معلومات صرفية، وإملائية، وصوتية، ونحوية، ودلالية لإزالة اللبس. ومن هنا أصبح تصميم المعجم وتطويره من القضايا الرئيسية في المعالجة الآلية للغات الطبيعية.(١)

إن الأداء النهائي لنظم المعالجة الآلية يعتمد على كفاءة المعجم المرتبط بها؛ إذ تؤدي المعلومات الصرفية، والنحوية، والدلالية، والسياقية -التي يُزَوِّد المعجمُ

---

(1) Brisco, T. (1991). Lexical issues in Natural Language Processing. In Klein, E. & Veltman, F. (1991). Natural Language and Speech. Springer, Verlag.

النظامَ بها- دورًا مهمًّا في هذا الأداء، ومع ذلك قد تتنوع مستويات التفصيل، ونوعية المعلومات التي يحتوي عليها المعجم باختلاف نظام المعالجة.

ويبدو هذا التنوع واضحًا عندما نقارن بين بناء معجم يقتصر استخدامه - مثلًا- على التصنيف النحوي للكلمات *Parts of speech tagging*، وآخر يُستَخدم في الإعراب الآلي *Automatic syntactic parsing*، أو الترجمة الآلية، أو غير ذلك من نظم المعالجة. فعند بناء مصنف نحوي، يجب أن يكون النظام قادرًا على تعيين كل الصيغ الممكنة للكلمة في اللغة، وبالتالي يحتوي المعجم المرتبط بالمُصنِّف النحوي على الوحدات المعجمية، مع معلومات حول التصنيف النحوي لها. ولا يحتاج هذا النظام إلى معلومات دلالية، أو تصنيفات فرعية *subcategorizations*، إلا في حالة استخدام نتائج هذا المُصنِّف مداخلَ للمعرب *Parser*، أو لتطبيقات أخرى أكثر تعقيدًا، كالترجمة الآلية.(١)

وعلى الجانب الآخر تحتاج بعض نظم الفهم الآلي للغات الطبيعية -كاسترجاع المعلومات، وإجابة الأسئلة -*Question answering* إلى معلومات تتعلق بمعاني الكلمات؛ حتى تقوم بعمل الاستنتاجات المناسبة. وعلى هذا يحتاج المعجم المرتبط بهذه النظم إلى أن يعكس العلاقات الدلالية كالترادف، والتقابل، والعلاقات الهرمية *Hierarchical relations* بين الكلمات؛ كي تقوم بتحليل «معنى» الجملة.

بناء المعجم: الجذر والوزن في مقابل الجذع:

تُمَثِّل الإجابة عن السؤال المتعلق بالأساس الذي يجب بناء المعجم الحاسوبي العربي عليه أهمية كبيرة قبل الشروع في بناء المعجم ذاته. وترتبط الإجابة عن هذا السؤال ارتباطًا وثيقًا بطبيعة اللغة العربية بوصفها لغة سامية من جهة، والاتجاه المُتَّبع في التحليل الصرفي لها من جهة أخرى؛ نتيجة للعلاقة الصرفية المعجمية التي تتسم بها هذه اللغات، ولأن اختيار المداخل المعجمية قائم على أسس صرفية بحتة. ويحاول هذا المبحث الإجابة عن هذا السؤال من خلال

---

(1) Megerdoomian, K. (2003). In Farghaly, A. (ed). Handbook for Language Engineers. Center for the Study of language and Information. P. 25.

تحليل اتجاهات الدرس الصرفي الحاسوبي، وتتبُّع فرضياتها وتأثيرها في بناء المعجم من الزاويتين: اللغوية، والحاسوبية.

عَدَّت الدراسات الأولية للصرف العربي الجذر الثلاثي مع وزن الكلمة الوحدتين الصرفيتين اللتين تتحكمان في بنية اللغة، فكلمة «كَتَبَ *kataba*» هي الشكل السطحي الذي يتحلل إلى الجذر الثلاثي (ك ت ب *ktb*)، والوزن (فَعَلَ *facafa*).

فالجذور -تبعًا لهذا الاتجاه في التحليل الصرفي- هي الأصول اللغوية للكلمات؛ أي الحروف الأصلية التي تُشتَق منها الكلمات. ومعظم الكلمات العربية ثلاثية الأصل، وهناك عدد قليل رباعي، أو خماسي. ويمكن القول بأن جذر الكلمة هو حروفها مجردة من الزوائد، وغير مقترنة بصيغة. ويحمل الجذر المعنى الأساسي للأنواع المعجمية المشتقة منه؛ فالجذر الثلاثي (ك ت ب) يحمل المعنى الأساسي للكتابة، وتشترك في هذا المعنى مشتقاته المختلفة، مثل: كَاتِب، وكِتَابَة، وكُتُب، ومَكْتُوب... إلخ.

وقد اتخذت معظم المعاجم العربية منذ بداية تأليفها جذر الكلمة أساسًا تورد تحته كافة أنواع المشتقات؛ اعتمادًا على هذين المبدأين الأساسيين:

- ❖ الجذر هو الأساس الذي تُشتق منه الكلمة.
- ❖ المشتقات المختلفة للجذر الواحد تشترك -ولو جزئيًا- في معنى أساسي يحمله الجذر.

وعلى المستوى الحاسوبي نالت دراسات الصرف العربي اهتمامًا كبيرًا من جانب مهندسي اللغة، واللغويين الحاسوبيين منذ بداية الثمانينات. وتفاعلت هذه الدراسات مع الجانب المعجمي تفاعلًا كبيرًا؛ إذ اقترحت باكورة الأعمال الخاصة بالصرف الحاسوبي العربي أن تكون فكرة الجذور والأوزان أساس تحليل الكتابة العربية المشكلة تشكيلًا كاملًا.(١)

---

(1) Dichy, J. & Farghaly, A. (2007). Grammar-Lexis Relation in The Computational Morphology of Arabic. In Soudi, A., Van Den Bosch, A. & Neumann, G. (eds.) (2007). Arabic Computational Morphology: Knowledge-based and Empirical Methods. Springer. P. 117.

وقد رأى أصحاب هذا الاتجاه أن طبيعة اللغة العربية تجعل من الصرف النظامي *Templatic morphology* المعتمد على النظرية العروضية، أي على فكرة الأوزان الصرفية، أكثر قدرة على التعامل مع الطبيعة غير التسلسلية للصرف في اللغة العربية، وفي غيرها من اللغات التي تتفق معها في هذه الطبيعة الصرفية.

واعتمادًا على هذا الاتجاه في التحليل الصرفي صُمِّم عدد من النظم الحاسوبية لتحليل الكلمات العربية أو توليدها. وتمثلت الطريقة الأساسية لبناء هذه النظم في عمل قاموس للجذور العربية، وقواميس للواحق، مع مراعاة التمييز بين السوابق واللواحق والدواخل، أو عمل معجم للجذور والأوزان تضاف إليه قوائم بالعناصر السابقة أو اللاحقة.(١)

وقد وجِّهت مجموعة من الانتقادات لهذا الاتجاه على أساس وجود بعض جوانب القصور، ومن ثم كانت الدعوة إلى تبني اتجاه بديل عند بناء قواعد البيانات المعجمية الحاسوبية. ولعل من أبرز الانتقادات التي وجِّهت إلى هذا الاتجاه:

- ❖ تمثيل الجذر والصيغة يصلح مع مجموعة فرعية من المعجم، في حين لا يمكن تطبيق الفكرة ذاتها على مجموعات أخرى أساسية:(٢)

لا يمكن تطبيق فكرة الجذر والوزن على الكلمات المعرَّبة، مثل: (بستان – فسفات– منجنيز– نيلون– سكرتير– بورصة– بنك– سندس– إستبرق– بلطجي)، كما لا يمكن تطبيقها على كثير من أسماء الأعلام (إشبيلية– غرناطة– دمشق– الفسطاط – إسماعيل– سيبويه).

---

(1) Dichy, J. & Farghaly, A. (2007). P. 120.

(2 Dichy, J. & Farghaly, A. (2003). Roots & Patterns vs. Stems plus Grammar-Lexis Specifications: on what basis should a multilingual lexical database centred on Arabic be built? MT Summit IX - Workshop: Machine Translation for Semitic Languages. New Orland, USA.

❖ تمثيل الجذر والصيغة يصلح -أساسًا- مع الأفعال والمشتقات الفعلية:[1]

تسري العلاقات الاشتقاقية من صيغة إلى أخرى -بشكل أساسي- على نطاق الأفعال والمشتقات الفعلية، كالمصدر، واسم الفاعل، واسم المفعول، ومن ثم يمكن تحليل الأفعال العربية، والمشتقات الفعلية على أساس الجذر والوزن، لكن هذا الاتجاه يترك عددًا كبيرًا من المداخل المعجمية دون تمثيل.

❖ عدم انتظام الفرضية القائلة بأن المشتقات المختلفة للجذر الواحد تشترك في المعنى نفسه:[2]

يدلل ديشي وفرغلي على عدم انتظام هذه الفرضية بأحد أشهر الأمثلة المستخدمة في الدراسات الصرفية، وهو الجذر (ك ت ب)؛ إذ يفترض أن المشتقات المعجمية المختلفة التي تحتوي على هذا الأصل الثلاثي تشترك في الحقل الدلالي للكتابة، على النحو الذي نراه في كُتُب، وكاتِب، ومَكْتَب، ومَكْتَبَة، وهو ما لا يبدو واضحًا -لمتحدثي العربية في الوقت الحالي- عند اشتقاق كلمة نحو «كتيبة»، التي يحتاج ربطها الدلالي بالجذر إلى دراسة تاريخية معنية بتطور اللغة وتغيرها عبر التاريخ.

وقد حاولت الدراسة تتبع هذه الفرضية من خلال البحث في المشتقات المعجمية لعدد من الجذور المختارة بطريقة عشوائية، وانتهت إلى أن اشتراك بعض الكلمات في حروف المادة له دلالته على الاتصال الشكلي بين هذه الكلمات، دون أن يعني -بالضرورة- اتصالاً بينها على مستوى الدلالة. كما يظهر من الأمثلة التالية:

- الجذر اللغوي (ر ق ب) يرتبط -أساسًا- بمعنى المراقبة، ومنه الفعل (راقَب)، واسم الفاعل (مُراقِب) بضم الميم وكسر القاف، واسم المفعول (مُراقَب) بضم الميم وفتح القاف، والمصدر (مُراقَبَة)، لكننا نحصل من الجذر ذاته على كلمة (رَقَبَة) التي تعني مؤخر أصل العنق.
- الجذر اللغوي (ش ك و) يرتبط -أساسًا- بمعنى الشكوى، ومنه الفعل شكا،

---

(1) Ibid.

(2) Ibid.

واسم الفاعل الشاكي، واسم المفعول مَشكُو، والمصدر شكاية، والاسم شكوى، لكننا نحصل من الجذر ذاته على كلمة (مِشْكاة) التي تعني الكُوَّة التي ليست بنافذة، وفي القرآن الكريم ﴿الله نور السموات والأرض مَثَلُ نوره كمشكاة فيها مصباح﴾[1] والمشكاة: موضع الفتيل من القنديل.

- الجذر اللغوي (ب ل د) نحصل منه على كلمة (البلد) بمعنى الموطن أو المُسْتَقر، ونحصل منه أيضًا على كلمة (البَلادَة) وهي نقيض الذكاء، ولا يوجد التقاء بينهما في الدلالة.

❖ الشذوذ وعدم الانتظام الاشتقاقي:

كثيرًا ما تترد في كتب الصرف جملة «وما جاء مخالفًا للقواعد شاذ، يُحفظ، ولا يُقاس عليه،» كما في الأمثلة التالية:[2]

- القاعدة: يُشتق اسما الزمان والمكان من الثلاثي على وزن (مَفْعَل) بفتح الميم والعين إذا كان مضارعه مضموم العين، نحو: (أَكَل - يَأْكُل - مَأْكَل)، و(بَلَغ - يَبْلُغ - مَبْلَغ)، لكننا نجد (مَسْجِد) بكسر العين من الثلاثي (سَجَد - يَسْجُد)، ونحو ذلك (مَطْلِع) من الثلاثي (طَلَع - يَطْلُع).
- القاعدة: يُشترط في اشتقاق اسم التفضيل على وزن (أفعل) أن يكون له فعل، ولكننا نجد: (هذا البعير أحنك الإبل)، أي أكثرها أكلًا بحكنه؛ فقد اشتق اسم التفضيل من (الحنك)، والحنك اسم، وليس فعلًا. ومن ذلك أيضًا (سعد ألصُّ من غيره)، أي أكثر منهم لصوصية؛ فقد اشتق اسم التفضيل من (اللص)، وليس بفعل، وإنما هو اسم.
- تشيع بصورة كبيرة المصادر غير القياسية، لا سيما مصادر الثلاثي حتى شاع الرأي بين الصرفيين على أنها سماعية، أي قد نجد أوزانًا للمصادر على غير قياس من ناحية، وقد نجد للوزن الواحد من الأفعال مصادر متعددة الصيغ من ناحية أخرى. فمن أوزان المصادر غير القياسية:

---

(١) سورة النور، الآية ٣٥.

(٢) راجع: أبو المكارم، علي. (١٩٨٠). تصريف الأسماء. الطبعة الأولى. المكتبة النحوية. ص ٦٤ - ٩٠.

❁ (فُعَل) بضم الفاء وفتح العين، نحو: هدى.

❁ (فَعْلة) بفتح الفاء وسكون العين، نحو: رحمة.

❁ (فَعَل) بفتح الفاء والعين معًا، نحو: طلب.

ومن قبيل تعدد صيغ المصادر للأفعال ذات الوزن الواحد مصادر الأفعال الدالة على الأصوات، نحو: صرخ، وعوى، وبكى، ونهق، وقد ورد فيها أوزان ثلاثة، هي:

❖ (فُعَال) بضم الفاء وفتح العين، نحو: صراخ، وبكاء.

❖ (فِعَال) بكسر الفاء وفتح العين، نحو: زمار (صوت النعام).

❖ (فَعِيل) بفتح الفاء وكسر العين، نحو: نهيق.

❖ اختلاف المعجميين واللغويين حول الجذر اللغوي لبعض الكلمات:

توجد كلمات عربية لا يقطع المعجميون بجذرها الحقيقي؛ فكلمة (ميناء) – مثلًا– تدرجها معاجم تحت الجذر (و ن ى)، وأخرى تحت الجذر(م ن أ)، وثالثة تحت الجذر (م أ ن).

على نحو مختلف، يرى فرغلي أن التوصيف الدقيق للصرف العربي يجب أن يميز بين ثلاثة مستويات: أولها الجذر *Root*، وهو غير منطوق وغير مصنف على مستوى الأنواع المعجمية، أو أنواع الكلام، والثاني الجذع *Stem*، وهو منطوق، ويجب أن يكون أحد الأنواع المعجمية في اللغة، والثالث الكلمة المتصرفة *Inflected words*؛ إذ تتصل بها الزوائد لتكوين معظم الكلمات العربية الفعلية،[1] كما في المثال الموضح بالشكل التالي.

المستوى الأول ◄— الجذر ◄— ك ت ب (غير منطوق – غير مصنف)

المستوى الثاني ◄— الجذع ◄— كَتَبَ (منطوق – مصنف)

المستوى الثالث ◄— الكلمات المتصرفة ◄— يكتب – يكتبون – تكتبون....

الشكل رقم (١)

المستويات الثلاثة للصرف عند علي فرغلي

---

(1) Farghaly, A. (1987). Three Level Morphology. Paper presented at the Arabic Morphology Workshop, Linguistic Summer Institute, Stanford, CA.

فـالجـذع الصـورة التـي يأخـذهـا الجـذر حـين يُفَـرَّغ فـي قـالـب من القـوالـب الصـرفـيـة، ومن ثم تتصل به السـوابق واللواحق لتكوين الصيغ المخـتلفـة للكلمـة. ويفـاد من ذلك أن ثمـة فـارقًا بـين الجـذر والجـذع يقـوم عـلى اعـتبـار الجـذر مـادة صوتية خـامًا، في حين يُعَدّ الجذع تحققًا من تحققات الجذر، كمـا يفاد أن الجذر الواحـد قـد يكون له غـيـر جـذع واحـد. وعـلى الرغم من اسـتـمـرار اعـتـمـاد بعض الدراسـات على فكرة الجذر والوزن في مـعـالجـة الصـرف العـربـي، فـإن الاتجـاه الحاسوبي الذي اقترحته معظم الدراسات الحديثة حول التعامل الصرفي للوحدة المعجمية-للغة العربية- يعد الجذع فحسب هو المناسب؛ إذ تتصرف على أساسه قواعـد التـحـقق. وقـد اعـتـمـد المعـجـم الذي طوره «تيم بكوالـتـر» ليكون جـزءًا من محلله الصرفي على الجذع؛ إذ يحتوي المعجم على قائمة من الجذوع الصرفية، مصنفة لأغراض التحليل التصريفي، ومصحوبة بالوحدات المعجمية العربية، والمقابلات الإنجليزية.

ويعـد مـعـجـم «تيم بكوالـتـر» من أشـهـر المعـاجم الحـاسـوبيـة، ولعله الأكـثـر استخدامًا في التطبيقات المختلفة لمعالجة اللغة العربية آليًا؛ فقد اعتمد عليه حبش ورامبو فى نظامهما المعروف باسم *MAGEAD*، كما اعتمد عليه أوتكار سمارتش في «الصرف العربي الوظيفي» بوصفه مصدرًا معجميًا أساسيًا.(١)

---

(١) انظر:

- *Habash, N & Rambow, O. (2006). MAGEAD: A Morphological Analyzer and Generator for Arabic Dialects. In Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meetings of the Association for Computational Linguistics. pp. 681 - 688. Association for Computational Linguistics, Australia: Sydne*

- *SmrZ, O. (2007). Functional Arabic Morphology: Formal System and Implementation. PhD Thesis, Charles University in Prague, Prague, Czech Republic.*

معجم قائم على أساس الجذع:

تقترح الدراسة اعتماد الجذع أساسًا لبناء المعجم على نحو يتجاوز جوانب القصور في المعاجم المعتمدة على فكرة الجذور والأوزان -من جهة- ويفيد من مميزات بناء المعجم على أساس الجذع من جهة أخرى، وهي مميزات تجعله أكثر فاعلية وسهولة في التطوير والتوسيع. ومن أبرز مسوغات بناء المعجم على أساس الجذع:

- ❖ التخلص من عملية توليد الجذوع من الجذور والأوزان الصرفية؛ فالمحللات الصرفية للغة العربية التي يعتمد معجمها على فكرة الجذر والوزن تحتاج إلى إدخال الجذور والأوزان الصرفية بوصفها معلومات ضرورية لتوليد الجذوع.[1] ولعملية التجذيع أهميتها في بعض التطبيقات، حين لا تكون هناك حاجة إلى تحليل الكلمة، ولكن إلى التوصل من خلال الكلمة إلى الجذع؛ ففي استرجاع المعلومات، والبحث على شبكة المعلومات قد نحتاج إلى التوصل من خلال كلمة «مسلمون» إلى الجذع «مسلم» دون الحاجة إلى معرفة أن كلمة «مسلمون» هي صيغة الجمع المتكونة بإضافة مورفيم جمع المذكر السالم «ون». ويطلق على هذا التطبيق «التجذيع».
- ❖ تُمَثِّل كل المواد المعجمية وحدات معجمية فعلية، وليس وحدات افتراضية.[2]
- ❖ يعد الجذع نوعًا معجميًا يمكن ربطه بالمعلومات اللغوية على مستويات الصرف، والنحو، والدلالة مثل إطارات التصنيف الفرعي *Sub-categorization frames*، وبنية المحمول *Argument Structure*، وذلك بخلاف الجذر غير المصنف على مستوى الأنواع المعجمية.
- ❖ الاعتماد على الجذع في بناء المعجم يساعد على التخلص من الصيغ غير المستعملة، والاقتصار على ما هو مستخدم فعليًا في مدونات اللغة العربية المعاصرة. ولهذه النقطة أهمية خاصة تتعلق بمفهوم شائع في الدراسات اللغوية

---

(1) Dichy, J. & Farghaly, A. (2003).

(2) Ibid.

والمعجمية يُعْرَف بالفجوات المعجمية[1] *Lexical gaps*؛ فكما يسمح نظام الجذور والأوزان بإنتاج الصيغ اللغوية المستعملة الصحيحة، قد يؤدي في الوقت ذاته إلى إنتاج صيغ مهملة، أو غير مستخدمة -فعليًا- عند متحدثي اللغة.

وتبدو فكرة الفجوات اللغوية واضحة في الجدول رقم (١) الذي يوضح المستعمل والمهمل من صيغ الجذر اللغوي (هـ ب ط).[2]

الجدول رقم (١)

المستعمل والمهمل من صيغ الجذر اللغوي (هـ ب ط).

الجذر	الصيغة	الكلمة	الاستعمال
هـ ب ط	(١) فَعَل	هَبَط	مستعمل
	(٢) فَعَّل	هَبَّط	مستعمل
	(٣) فَاعَل	هَابَط	غير مستعمل
	(٤) أَفعَل	أَهْبَط	مستعمل
	(٥) تَفَعَّل	تَهَبَّط	مستعمل
	(٦) تَفَاعَل	تَهَابَط	غير مستعمل
	(٧) انْفَعَل	انْهَبَط	مستعمل
	(٨) افْتَعَل	اهْتَبَط	غير مستعمل
	(٩) افْعَلّ	اهْبّطّ	غير مستعمل
	(١٠) اسْتَفْعَلَ	اسْتَهْبَط	غير مستعمل
	(١١) افعوعل	اهْبَوْبَط	غير مستعمل

---

(١) يستخدم مصطلح الفجوة في الدراسات اللغوية -بشكل عام- للإشارة إلى غياب وحدة لغوية ما في قالب من العلاقات يكون وجود هذه الوحدة اللغوية فيه متوقعًا. وقد تظهر هذه الفجوات في مستويات الدرس اللغوي المختلفة؛ فهناك الفجوات الفونولوجية *Phonological gaps*، والفجوات الصرفية *Morphological gaps*، والفجوات النحوية *Syntactic gaps*، والفجوات المعجمية *Lexical gaps*، انظر:

- *Crystal, D. (2008). A Dictionary of Linguistics and Phonetics. Sixth edition. Blackwell Publishing Ltd. P. 205.*

(٢) اعتمدت الدراسة في تحديد المستعمل من الصيغ الفعلية للجذر (هـ - ب - ط) =

تنقسم الجذوع في اللغة العربية -تبعًا لديشي وحسون- إلى نوعين عامّين.[1] وهو التقسيم القائم على إمكانية إخضاع الجذع للتحليل تبعًا لفكرة الجذور والأوزان، أو عدم خضوع الجذع لهذه الفكرة.

- ❖ النوع الأول: يشتمل على الأنواع المعجمية الأساسية التي يمكن تمثيلها باعتبار الجذر والوزن، وتنتمي الأفعال والمشتقات الفعلية إلى هذا النوع، كما في الجذع اللغوي «تَكَبَّر» الذي يتألف من الجذر الثلاثي (ك ب ر)، والوزن (تَفَعَّل).
- ❖ النوع الثاني: يطلق ديشي على هذا النوع أشباه الجذوع *quasi-stems*، ولا يشتمل -بطبيعة الحال- على الأفعال والمشتقات الفعلية، ولكنه يشتمل على بعض الأسماء التي لا يمكن تحليلها باعتبار الجذور والأوزان، كما في الجذر اللغوي «بَرْنَامَج»؛ إذ لا يمكن اشتقاق أي فعل، ولا جمع مذكر سالم من الصوامت الخمسة (ب ر ن م ج).

الخصائص الصرفية النحوية والدلالية التي تُرَمَّز في المعجم:

مع ظهور اللغويات الحاسوبية، وبالنظر إلى المعلومات التي تتطلبها نظم المعالجة الآلية، وإلى الطبيعة التصريفية للغة العربية، لم يعد التصنيف التقليدي للكلمة إلى اسم، وفعل، وحرف كافيًا للدلالة على بنية الكلمة؛ فالعربية تشتمل على خصائص صرفية ونحوية عديدة، كالتقسيمات الفرعية، والعدد، والشخص، والحالة الإعرابية... إلخ، ولذلك «لا يُستخدم هذا التصنيف -غالبًا- في الدراسات الحاسوبية،»[2] وإنما تستخدم تصنيفات أخرى أكثر تفريعًا تراعي

= على ثلاثة معاجم: لسان العرب، وتاج العروس باعتبارهما الأغزر مادة، والمعجم الوسيط لاحتمال استعمال العربية الحديثة صيغًا كانت مهملة.

(1) Dichy, J. & Hassoun, M. (eds). (1989). Simulation de modèles linguistiques et Enseignement Assiisté par Ordinateur de l arabe - Travaux SAMIA I. Paris, Conseil International de la langue Francaise. Depending on: Dichy, J. & Farghaly, A. (2003 - 2007).

(2) Habash, N. (2010). Introduction to Arabic Language Processing. Morgan & Claypool Publishers Series. P. 79.

الخصائص الصرفية، والنحوية، والدلالية التي تستخدمها اللغة. وقد طُوِّرت بالفعل تصنيفات عدة للكلمة العربية منها تصنيف مدونة بنسلفانيا العربية *PATB tag set*، وتصنيف شيرين خوجة *Khoja tag set*، والتصنيف الصرفي لمدونة القرآن الكريم *The Quranic Arabic Corpus morphological tag set*، وتصنيف مجدي صوالحه الذي استخدمه في محلله الصرفي الذي أطلق عليه *SALMA tag set*، وغيرها. ولأن الدراسات اللغوية الحاسوبية كانت أسبق عند الأوروبيين، فقد سبق هذه التصنيفات العربية مشروع المجموعة الاستشارية الخبيرة في معايير هندسة اللغة «*EAGLES*»، وهو المشروع الذي يهدف إلى اقتراح مجموعة الخصائص الصرفية النحوية العامة التي تُرَمَّز في المعاجم والمدونات النصية للغات الأوروبية.

التصنيف المقترح للأنواع المعجمية:

ترتبط أهمية تصنيف الأنواع المعجمية لتطبيقات المعالجة الآلية للغات الطبيعية بكم المعلومات الكبير الذي يعطيه التصنيف عن الكلمة[1] . وتظهر أهمية تصنيف الكلام في استرجاع المعلومات -مثلا- في معرفة سوابق الكلمة ولواحقها، مما يساعد في عملية التجذيع، فالتاء المربوطة -مثلًا- تلحق الاسم والصفة ولا تلحق الفعل، وكذلك مورفيم جمع التأنيث (ات)، في حين تحلق التاء المفتوحة الفعل، ولا تلحق الاسم.

تنقسم الكلمات إجمالاً إلى قسمين: مجموعة مفتوحة *Open category* ومجموعة مغلقة *Closed category*. ويقوم هذا التقسيم على وجود أوجه اختلاف عدة بين القسمين، منها الحجم، والعضوية، والمعنى، ومعدل التكرار في النصوص، على النحو الموضح بالجدول التالي:[2]

---

(1) Jurafsky, D. & Martin, J. (2006). Speech and Language Processing: An Introduction to Language Processing. Ch. 5. P.2

(2) Haspelmath, M. (2001). Word Class and Parts of Speech. In the International Encyclopedia of Social & Behavioral Sciences. Elsevier Science Ltd.

الجدول رقم (٢)

أوجه الاختلاف بين المجموعة المفتوحة والمغلقة

وجه الاختلاف	المجموعة المفتوحة	المجموعة المغلقة
الحجم	كبير	صغير
العضوية	مفتوحة «تقبل كلمات جديدة»	مغلقة «لا تقبل كلمات جديدة»
المعنى	ملموس، ومحدد	مجرد، وعام (أو بدون معنى)
معدل التكرار	منخفض	عال

وعلى هذا تنتمي حروف الجر «من» و«إلى» و«على» إلى المجموعة المغلقة؛ لأنها مجموعة ثابتة من الكلمات، وصغيرة الحجم، وليس لها معنى في ذاتها وإنما تكتسب معناها من دخولها على الاسم، كما أنها شديدة الشيوع في النصوص. والأمر ذاته ينطبق على كل الكلمات الوظيفية *Function words*، وهي كلمات دورها الأساسي التعبير عن علاقات نحوية[١] وفي المقابل تدخل الأسماء ضمن المجموعة المفتوحة؛ لأنها غير ثابتة، بل تتجدد دائما وتدخل أسماء جديدة إلى اللغة بالاستعارة من لغات أخرى، أو بالتوسع في القياس، كما أنها كبيرة الحجم، وهي أقل شيوعًا واستخدامًا داخل النصوص من الكلمات الوظيفية. والأمر ذاته يسري على باقي كلمات المحتوى -وهي الكلمات التي لها معنى معجميٌّ يمكن التعبير عنه[٢] -، وهي الأسماء، الأفعال، والصفات، والظروف.

المجموعة المفتوحة:

تشتمل المجموعة المفتوحة على أربعة أنواع أساسية: الاسم، والفعل، والصفة، والظرف. ويعتمد اختيار النوع المعجمي لكلمات هذه المجموعة على ثلاثة معايير: معيار مفاهيمي، وآخر صرفي، وثالث توزيعي.

---

(1) Crystal, D. (2008). P. 108.

(2) Ibid.

الأسماء:

تقليديًا، صُنِّفت الكلمات التي تشير إلى أسماء الأشخاص، والأماكن، والأشياء باعتبارها «أسماءً»، وهذا التصنيف له أهميته في دلالة الاسم على مسمى، سواء أكان هذا المسمى شخصًا أم شيئًا، لكن الغموض المرتبط بمفاهيم مثل «الأشياء»، كالسؤال حول ماهية «الجمال» دفع إلى تحليل هذا النوع المعجمي باعتبارات أخرى صرفية وتوزيعية.(١) ويمكننا أن نطبق المعايير الثلاثة لتحديد الكلمات التي تنتمي إلى الأسماء:

مفاهيميًا: تشير الأسماء بوجه عام إلى الكائنات الحية، مثل «أسد»، «غلام»، و«شجرة»، وإلى الأماكن كـ«القاهرة»، و«جدة»، «الكويت»، وإلى الأشياء مثل «كتاب»، و«حذاء»، و«باب»، وإلى المفاهيم كـ«الحب»، و«السعادة»، و«الزواج».

صرفيًا: حدد النحاة العرب معيارين لتحديد الاسم في اللغة العربية، هما: قبول أداة التعريف «أل»، كما في «الكتاب»، و«الغلام»، والشجرة. وقبول التنوين في نحو «كتابٌ»، و«غلامٌ»، و«شجرةٌ». كما أن الأسماء العربية تكشف عن أنواع من التصريف كالعدد، فيقال في المفرد «كتاب»، و«غلام»، و«شجرة»، ويقال في المثنى «كتابان»، و«غلامان»، و«شجرتان»، وفي النوع يقال «طالب»، و«طبيب» للمذكر، و«طالبة»، و«طبيبة» للمؤنث. وفي الحالين يقبل الاسم مورفيمًا تصريفيًا.

توزيعيًا: قد تأتي الأسماء بعد حروف الجر في نحو «في الكتاب»، و«مع الغلام»، و«على الشجرة»، كما تؤدي وظيفة نحوية محددة في الجملة كالفاعل والمفعول(٢) في نحو «الكتاب خير صديق»، و«قرأت كتابًا»، كما تدخل عليها دون غيرها أداة النداء مثل «يا محمد»، و«يا فاطمة».

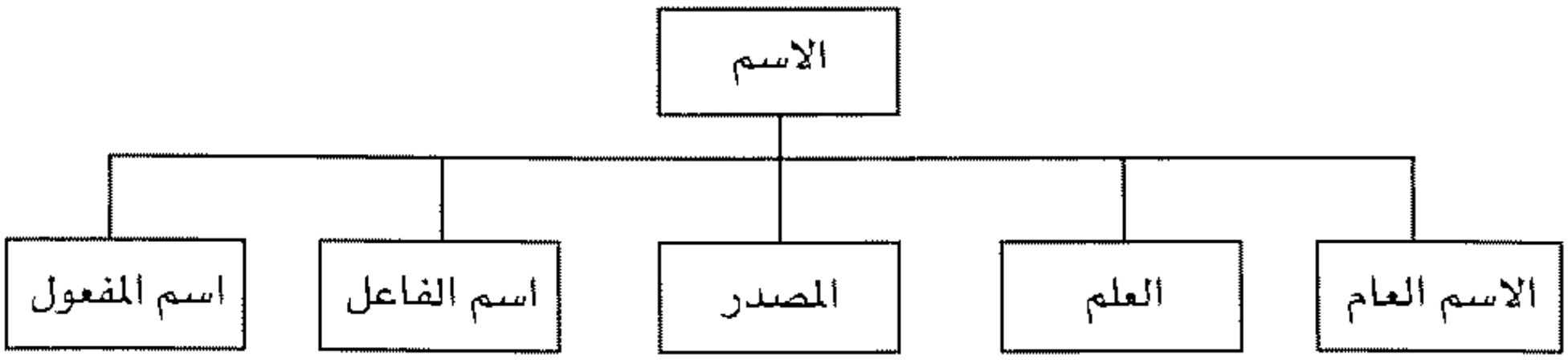
تنقسم الأسماء إلى خمسة أقسام: اسم عام *Common Noun*، وعلم *Proper Noun*، واسم فاعل *Active Participle*، واسم مفعول *Passive Participle*، ومصدر *Verbal Noun*. وينقسم الاسم العام -دلاليًا- إلى اسم ذات *Concrete Noun* واسم معنى *Abstract Noun*. كما يميز التصنيف على المستوى الدلالي بين الأسماء التي

---

(1) Crystal, D. (2008). P. 333.

(2) Ibid.

تشير إنسان *Human* مثل «رجل»، و«فتاة»، و«سفير»، و«مدير»، و«أستاذ»، في مقابل الأسماء التي لا تشير إلى إنسان *Inhuman* مثل «مسجد»، و«كتاب»، و«راية»... إلخ. يوضح الشكل رقم (٢) التقسيم الفرعي للأسماء.



الشكل رقم (٢)

التقسيم الفرعي للأسماء

أما أسماء الأعلام، فتضم الأنواع الموجودة في التقسيم الذي اقترحه ووجن بيك وآخرون، وهو التقسيم الذي يشتمل على تسع عقد فرعية، وثلاثين عقدة نهائية،[١] مع استثناء العقدة الفرعية الخاصة بالانتساب؛ لاعتبار النسبة من الصفات، وكذلك العقدة النهائية الخاصة بحكومة الولايات المتحدة، واعتبارها ضمن العقدة النهائية الخاصة بالحكومة، على النحو الموضح بالجدول رقم (٣).

---

(1) Paik, W., Liddy, E., Yu, E. & McKenna, M. (1993b). Categorizing and Standardizing Proper Nouns for Efficient Information Retrieval. In B. Boguraev & J. Pustejovsky (Eds.), Proceedings of the workshop on acquisition of lexical knowledge from text, Ohio, USA, pp.154-160. Cambridge, MA: MIT Press.

الجدول رقم (٣)

تقسيم ووجن بيك وآخرين لأسماء الأعلام

كيان جغرافي			منظمة	بشري	وثيقة	أدوات	علمي	زمني	مختلف
مدينة	ميناء	مطار	شركة	شخص	وثيقة	برمجيات	أمراض	تاريخ	مختلف
جزيرة	مقاطعة	حي	نوع شركة			أجهزة	عقاقير		
دولة	قارة	إقليم	حكومة	لقب		آلات	كيماويات	وقت	
مائي	مناطق أخرى		منظمة						

الخصائص التصريفية للأسماء:

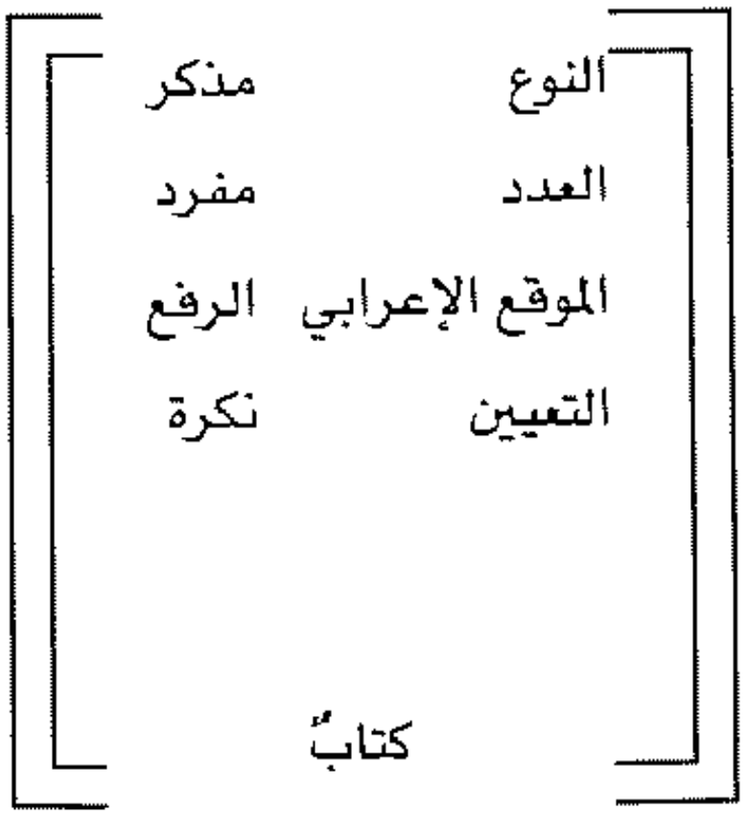
هناك حالات ثماني تُصَرَّف على أساسها الأنواع المعجمية العربية، هي: الزمن، والشخص، وبناء الفعل، والحالة الإعرابية للفعل، والنوع، والعدد، والموقع الإعرابي، والتعيين. وتختلف الأنواع المعجمية فيما بينها من حيث تصرفها أو عدمه تبعًا لكل حالة من هذه الحالات؛ فتُصرَّف الأسماء وفقًا لأربع حالات هي: النوع، والعدد، والموقع الإعرابي، والتعيين، على النحو الموضح بالجدول رقم (٤).

الجدول رقم (٤)

حالات تصريف الاسم

النوع		
مذكر	مؤنث	مذكر/مؤنث
العدد		
مفرد	مثنى	جمع	غير معدود
الموقع الإعرابي		
الرفع	النصب	الجر
التعيين	
نكرة	معرفة

وعلى هذا يكون التمثيل المعجمي للمحتوى الصرفي للأسماء مشتملًا على المعلومات الصرفية الخاصة بالنوع، والعدد، والموقع الإعرابي، والتعيين كما في التمثيل المعجمي للمحتوى الصرفي للاسم «كتابٌ»، الموضح بالشكل (٣).



الشكل رقم (٣)

التمثيل المعجمي للمحتوى الصرفي للاسم «كتابٌ»

الأفعال:

يشار إلى الأفعال بأنها الكلمات الدالة على الحدث والزمن جزء منها، فإذا كان من الأسماء ما يدل على الحدث مثل كلمة «كتابة»، و«ذهاب»، و«نوم»، فإن الفعل يختلف عن هذه الأسماء في دلالته مع الحدث على الزمن. ويمكننا تطبيق المعايير الثلاثة السابقة على الفعل:

مفاهيميًا: الفعل هو الكلمة الدالة على حدث مقترن بزمن؛ فالأفعال «كتب»، و«يكتب»، و«اكتب» تشير جميعها إلى حدث «الكتابة»، لكن الأول قيد الحدث بزمن المضي، في حين ربطه الثاني بالحال الذي لم ينقطع، أما الثالث فطلب فعله في زمن مستقبلي؛ أي أنه لم يقع بعد.

صرفيًا: ميز النحاة العرب الفعل بقبول تاء الفاعل وتاء التأنيث، وهو تمييز خاص بالفعل الماضي، مثل «كتبتُ» و«كتبتْ»، وميزوا المضارع بقبول نون التوكيد، ولام التوكيد مثل «لأكتبنَّ» وكذلك بدخول سين الاستقبال عليه، نحو «سيكتب».

ومُيِّز الأمر بقبول ياء المخاطبة، نحو «اكتبي».

يحدد ديفيد كريستال الفعل -من المنظور الشكلي- بأنه «العنصر الذي يمكن أن يكشف عن تباين صرفي في الزمن، والجهة، والبناء، والحالة الإعرابية، والشخص، والعدد».[1] ويمكن تطبيق ذلك على الفعل «كتب»:

الزمن	كتب (الماضي) - يكتب (المضارع)
الجهة	كتب (التام) - يكتب (غير التام)
البناء	كَتَبَ (المعلوم) - كُتِبَ (المجهول)
الحالة الإعرابية	يكتبُ (الرفع)- يكتبَ (النصب)
الشخص	أكتب (المتكلم) - اكتب (المخاطب) - يكتب (الغائب)
العدد	كتب (المفرد) - كتبا (المثنى) - كتبوا (الجمع)

ويزيد على ما حدده كريستال -بالنسبة للفعل في اللغة العربية- الجنس أو النوع؛ إذ تميز العربية بين المذكر والمؤنث، كما «كتب» و«كتبت»، و«يكتب» و«تكتب»، و«اكتب» و«اكتبي».

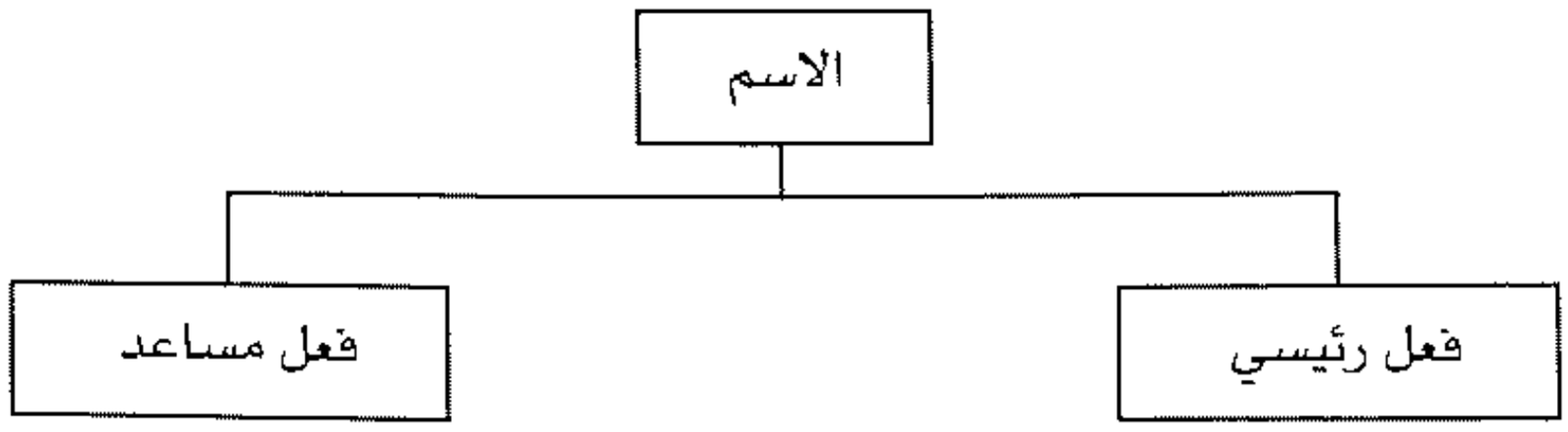
توزيعيًّا: تقوم الأفعال -دائما- بدور المسند *Predicate*، وتأخذ ما بين محمول واحد إلى أربعة محمولات. ومن هذه الجهة ينقسم الفعل إلى لازم، مثل «مات»، وهو الذي يأخذ محمولًا واحدًا، ومتعد، وهو الذي يقبل ما بين محمولين إلى أربعة، مثل «كتب»، و«أعطى»، ومن الأفعال التي تأخذ أربعة محمولات «أعلم» «أنبأ» ويطلق عليها في النحو العربي «الأفعال المتعدية لثلاثة مفاعيل» كما في جملة «أعلمت خالدًا محمدًا مسافرًا». قد تسبق بعضَ الأفعال في اللغة العربية أدواتٌ معينة كدخول «لن» و«لم» على الفعل المضارع في نحو «لن أكتب»، و«لم أكتب».

تنطبق هذه المعايير انطباقا تامًّا على أغلب الأفعال في اللغة العربية على النحو الذي رأيناه في الفعل «كتب»، وكذلك الأفعال «ضرب»، و«لعب»، و«قرأ»

---

(1) Crystal, D. (2008). P. 510.

وغيرها. وتنطبق بصورة جزئية على مجموعة أخرى من الأفعال ذات خصائص مميزة قد تعبر عن تنوعات في الزمن أو الجهة، مثل «كان وأخواتها». ولعل هذا ما دفع تمام حسان في دراسته عن العربية معناها ومبناها إلى التعامل مع هذه الأفعال باعتبارها أدوات، معللا ذلك بجمود بعضها وعدم تصرفه، وكذلك دخولها على الأفعال[1]. تتشابه هذه الأفعال في وظائفها مع الأفعال المساعدة في اللغة الإنجليزية، ولذلك تقسم الدراسة الأفعال إلى قسمين: أفعال رئيسية *Main verbs* وأفعال مساعدة *Auxiliary verbs*. تتناول الأفعال الرئيسية في هذا الجزء وتتوقف أمام الأفعال المساعدة عند الحديث عن المجموعة المغلقة.



الشكل رقم (٤)

التقسيم الفرعي للأفعال

الخصائص التصريفية للأفعال:

يصرف الفعل وفقًا لست حالات هي: الزمن، والشخص، وبناء الفعل، والحالة الإعرابية، والنوع، والعدد، كما في الجدول رقم (٥).

---

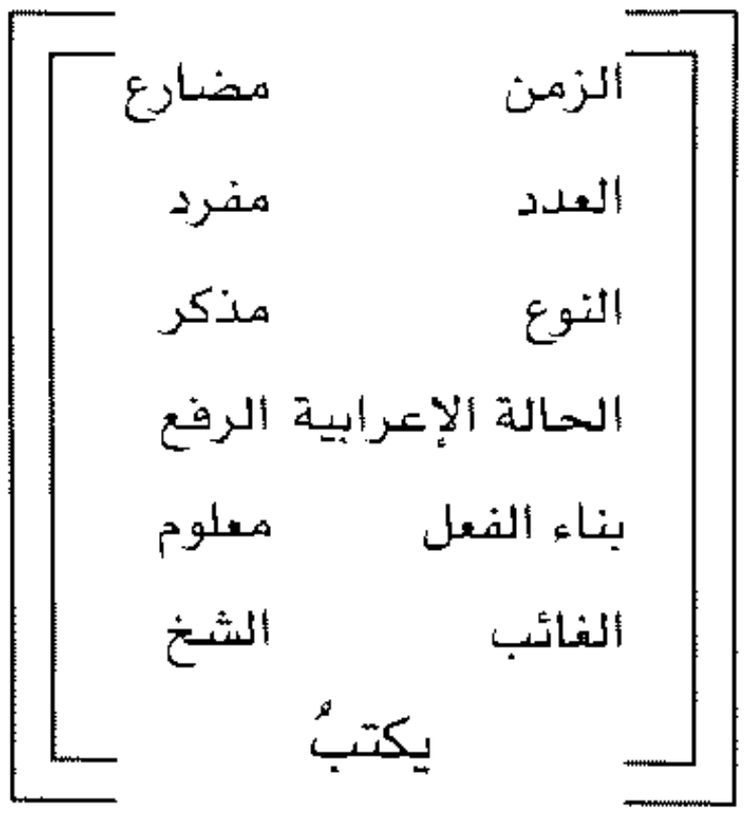
(١) حسان، تمام. (١٩٩٨). اللغة العربية: معناها ومبناها. الطبعة الثالثة. القاهرة: عالم الكتب. ص ١٢٩ - ١٣٢.

الجدول رقم (٥)

حالات تصريف الفعل

الزمن		
ماضٍ	مضارع	أمر
العدد		
مفرد	مثنى	جمع
الحالة الإعرابية		
الرفع	النصب	الجزم
النوع		
مذكر	مؤنث	
بناء الفعل		
بناء للمعلوم	بناء للمجهول	
الشخص		
المتكلم	المخاطب	الغائب

وعلى هذا يكون التمثيل المعجمي للمحتوى الصرفي للأفعال مشتملاً على المعلومات الصرفية الخاصة بالزمن، والعدد، والنوع، والحالة الإعرابية، وبناء الفعل، والشخص كما في التمثيل المعجمي للمحتوى الصرفي للفعل «يكتبُ»، الموضح بالشكل (٥).



الشكل رقم (٥)

التمثيل المعجمي للمحتوى الصرفي للفعل «يكتبُ»

تنقسم الأفعال من جهة المفعول إلى قسمين: الأفعال اللازمة، وهي الأفعال التي تكتفي بفاعلها، ولا تقع على مفعول به، وأفعال متعدية: منها ما يتعدى بنفسه إلى المفعول دون حاجة إلى حرف جر، ومنها ما يحتاج إلى حرف جر. والأفعال المتعدية تنقسم من حيث عدد مفاعيلها ثلاثة أقسام: أفعال تتعدى إلى مفعول واحد، وثانية تتعدى إلى مفعولين، وهي على ضربين: أفعال تتعدى إلى مفعولين أصلهما المبتدأ والخبر، وأخرى تتعدى إلى مفعولين ليس أصلهما المبتدأ والخبر، وثالثة تتعدى إلى ثلاثة مفاعيل، وبناء على هذا الاختلاف تميز الدراسة بين القسمين.

الصفات:

عدَّ النحاة العرب القدماء الصفة من الأسماء، وعدها بكوالتر قسمًا فرعيًا مما أطلق عليه الاسميات *Nominals*، في حين تتعامل الدراسة معها باعتبارها نوعًا معجميًا مستقلًا. ويمكننا أن نطبق المعايير الثلاثة لتحديد الكلمات التي تنتمي إلى الصفات:

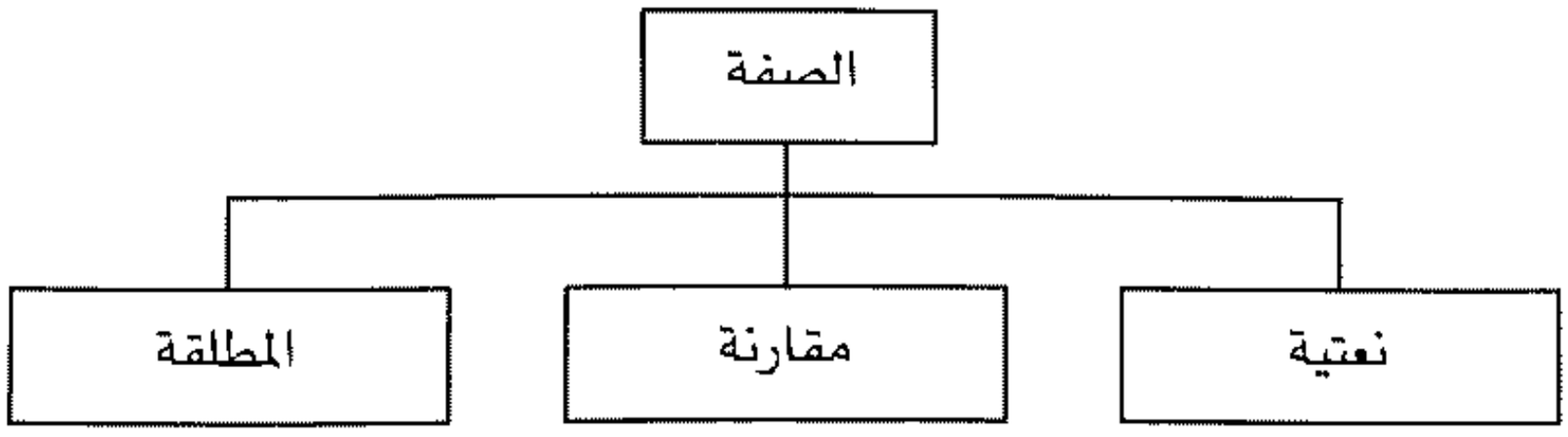
مفاهيميًا: الصفات هي مجموعة من الوحدات المعجمية التي تستخدم لتحديد خصائص الأسماء، فهي لا تدل على مسمى كالأسماء، وإنما تحدد خصائص هذه الأسماء.

صرفيًا: تتفق الصفة مع الاسم في التصريف من جهة النوع، والعدد، وكذلك في قبولها أداة التعريف «أل». كما يمكن استخدام الصفة في صيغ مقارنة نحو «أكبر»، و«أطول»، و«أجمل»، وقد يتعذر ذلك فتستخدم استراتيجية شبيهة باستخدام كلمة «*More*» في اللغة الإنجليزية؛ إذ تسبق كلمة مثل «أكثر»، أو «أقل»، أو «أشد» اسمًا نكرة في حالة النصب، كما في «أكثر عدوانية»، و«أشد تعاطفًا»، و«أقل تركيزًا». كذلك عندما تسبق أداة التعريف «أل» الصيغ المقارنة تحولها إلى صيغة التفضيل العليا، كما «أكبر» و«الأكبر»، و«أقدم» و«الأقدم»، و«أشهر» و«الأشهر».

توزيعيًا: تتبع الصفة الاسم مباشرة؛ أي أنها ترد في موقع الوصف داخل العبارة الاسمية، مثل «رجل كبير»، و«الفتاة الجملية»، و«شعر طويل». كما يمكن

المبالغة في معنى الصفة باستخدام كلمات مثل «جدًا»، نحو «كبير جدًا»، و«طويل جدًا»، و«جميل جدًا».

تميز الدراسة بين ثلاثة استخدامات للصفة: الوصف وتسمى الصفة في هذه الحالة الصفة النعتية *Attributive adjective*، وتستخدم له ما يعرف بالصفة المشبهة كـ«سهل»، و«صعب»، و«فرح»، و«فطن»، و«ضخم» و«حلو» ...إلخ، وصيغ المبالغة مثل «صبور، و«ظلوم»، و«جهول»، والنسبة مثل «مصري»، و«عراقي»، و«رأسمالي»، و«ناصري»، و«كنسي»، و«صحراوي». أما الاستخدام الثاني فهو المقارنة وتسمى الصفة في هذه الحالة الصفة المقارنة *Comparative adjective*، ويستخدم للمقارنة أفعل التفضيل، مثل «أسعد»، و«أفضل»، و«أقل»، و«أصغر»، وأخيرًا للدلالة على صيغة التفضيل العليا، ويتم ذلك في اللغة العربية بإضافة أداة التعريف «أل» إلى أفعل التفضيل، نحو «الأسعد»، و«الأفضل»، و«الأقل»، و«الأصغر». يوضح الشكل رقم (٦) التقسيم الفرعي للصفات.



الشكل رقم (٦)

التقسيم الفرعي الصفات

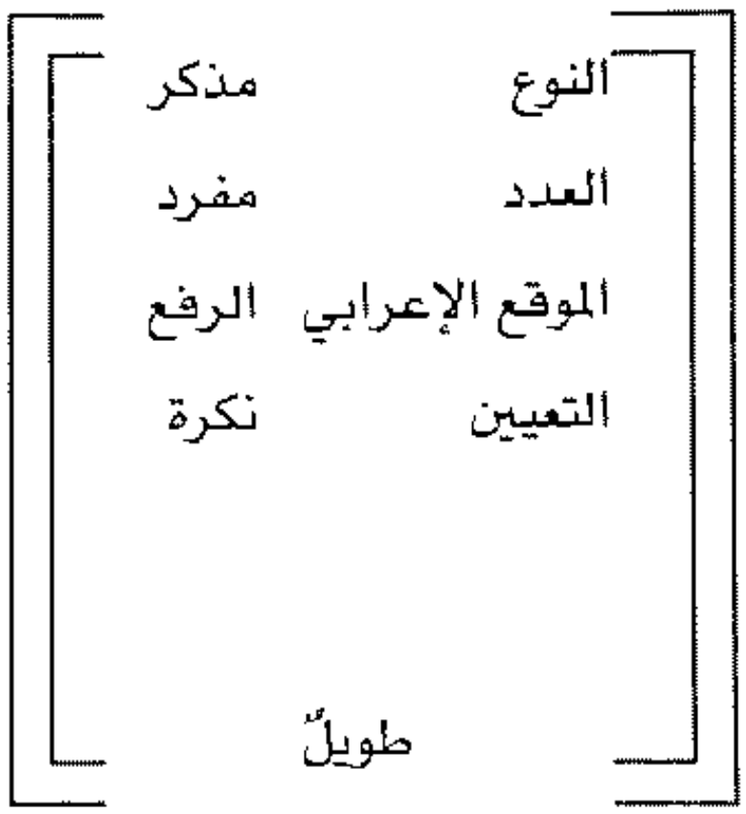
الخصائص التصريفية للصفات:

تتفق الصفات مع الأسماء في تصرفها تبعًا للنوع، والعدد، والموقع الإعرابي، والتعيين، على النحو الموضح بالجدول رقم (٦).

الجدول رقم (٦)
حالات تصريف الصفة

النوع		
مذكر	مؤنث	مذكر/مؤنث
العدد		
مفرد	مثنى	جمع
الموقع الإعرابي		
الرفع	النصب	الجر
التعيين		
نكرة	معرفة	

وعلى هذا يكون التمثيل المعجمي للمحتوى الصرفي للصفات مشتملًا على المعلومات الصرفية الخاصة بالنوع، والعدد، والموقع الإعرابي، والتعيين كما في التمثيل المعجمي للمحتوى الصرفي للصفة «طويلٌ»، الموضح بالشكل (٧).



الشكل رقم (٧)
التمثيل المعجمي للمحتوى الصرفي للصفة «طويلٌ»

الظروف:

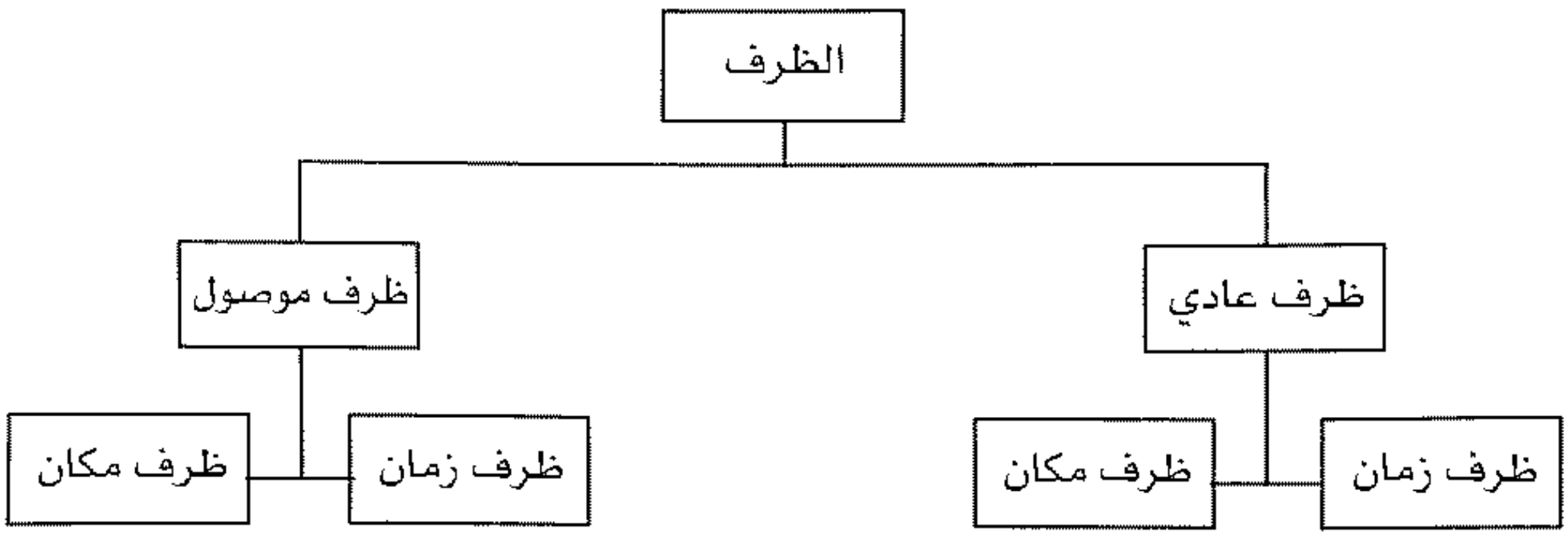
يتضح عند تطبيق المعايير الثلاثة المذكورة على الظرف أننا إزاء نوع معجمي يختلف عن الأنواع المعجمية السابقة، غير أن الكلمات التي تعد ظروفًا بذاتها قليلة في اللغة العربية مثل «قط»، و«إذ»، و«إذا»، و«أيّان»، و«حيث»، و«متى»، و«أين»، ولذلك تستعيض اللغة العربية عن ذلك باستخدام الأسماء أو الصفات في حالة النصب للدلالة على الظرفية، مثل «أحيانًا»، و«غدًا»، و«سريعًا»، و«قليلًا»، و«مبكرًا»، و«متأخرًا»، و«دائمًا».

مفاهيميًا: الظروف عبارة عن مجموعة من الوحدات المعجمية التي تضيف معلومات محددة حول زمن الفعل، مثل «أصلُ يوم السبت» أو مكانه نحو «جلس تحت الشجرة»، أو كيفيته، مثل«جاء مسرعًا».

صرفيًا: الظروف في اللغة العربية كلمات غير متصرفة، فهي تختلف عن الأسماء والأفعال والصفات من هذه الجهة.

توزيعيًا: قد يأتي الظرف بعد الفعل مباشرة، وقد يأتي الظرف في أول الجملة.

تنقسم الظروف إلى ظروف عادية *Regular adverbs*، وظروف موصولة *Relative adverbs*، وتتفرع هذه الظروف إلى ظروف للزمان *Time adverbs* وظروف للمكان *Location adverbs*. يوضح الشكل رقم (٨) التقسيم الفرعي للظروف.



الشكل رقم (٨)

التقسيم الفرعي للظروف

المجموعة المغلقة:

تشتمل المجموعة المغلقة على الأعداد، وكلمات الكم، والضمائر، والأفعال المساعدة، وكلمات الانفعال *Interjections*، والحروف. تتباين عناصر هذه المجموعة اقترابًا وابتعادًا من عناصر المجموعة المفتوحة؛ إذ تتشابه الأعداد والكلمات الدالة على الكم، والضمائر في جوانب عدة مع الأسماء والصفات، في حين تتشابه الأدوات مع الظروف.

الأعداد:

تنقسم الأعداد قسمين: أعداد أصلية *Cardinal numbers*، وهي الأعداد التي تدل على مقادير الأشياء المعدودة مثل «خمسة رجال»، و«سبع سيدات»، و«تسعة كتب»، كالأعداد التي تبدأ بـ«صفر» إلى العشرة، وكذلك ألفاظ العقود، وسالعدد «مائة»، و«مليون» ...إلخ. أما القسم الثاني فيضم الأعداد الدالة على رتب الأشياء المعدودة، ولذلك يطلق عليها الأعداد الترتيبية *Ordinal numbers*، مثل «الثاني»، و«الخامس»، و«التاسع».

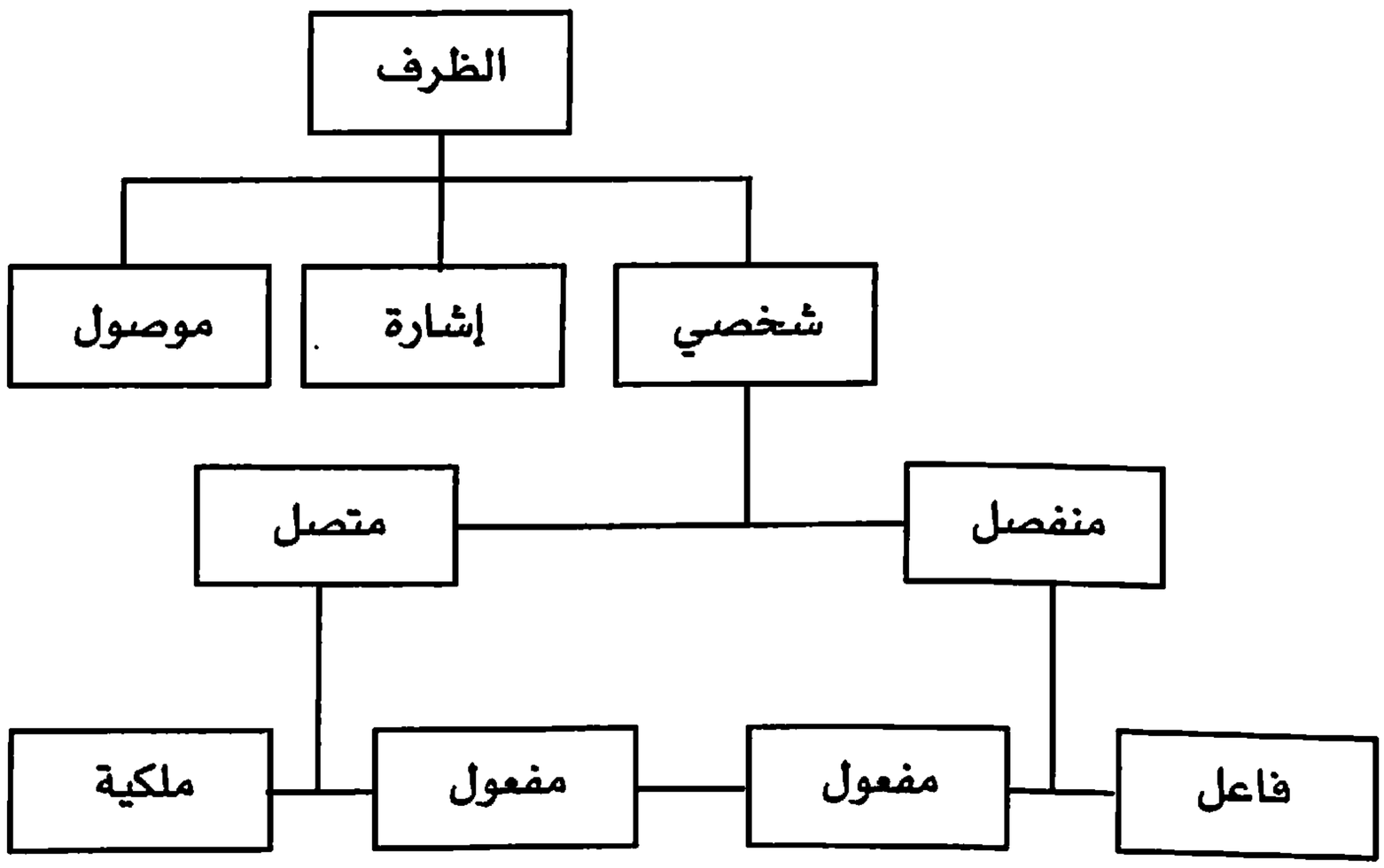
كلمات الكم:

يقصد بها الكلمات التي تعبِّر عن التباين في الكمِّ، ومن أشهرها كل، وبعض، وبِضعة، وشَطْر، وضِعف، وأغلب، وقُرابة... وغيرها.

الضمائر:

الضمائر هي الكلمات التي يمكن أن تحل محل الاسم، أو العبارة الاسمية، وتكون للإشارة إلى غائب، أو حاضر (متكلم أو مخاطب). تنقسم الضمائر إلى ثلاثة أقسام أساسية: ضمائر شخصية *Personal pronouns*، وهي تنقسم بدورها إلى ضمائر منفصلة، وأخرى متصلة. تنقسم الضمائر المنفصلة إلى ضمائرَ للفاعل، مثل «أنا»، و«أنت»، و«هو» و«هي»، وضمائر المفعول، مثل«إياي»، و«إياك»، و«إياه»، و«إياها». أما الضمائر المتصلة فتأتي مع الاسم والفعل وحرف الجر، لكنها تختلف -تمامًا- في دلالتها إذا لحقت الاسم عنها إذا لحقت الفعل أو

الحرف؛ فتعطي في الحالة الأولى معنى الملكية في نحو «كتابي»، و«قلمها»، و«معلمنا»؛ أي أنها تقابل الضمائر «*my*»، و«*her*» و«*our*» في العبارات «*my book*»، و«*her pen*»، و«*our teacher*» على الترتيب، ومن ثم تصنف باعتبارها ضمائر ملكية *Possessive pronouns*. أما إذا لحقت الفعل أو الحرف فتعطي معنى المفعولية كما في «ضربه»، و«معنا»، ومن ثم تقابل الضميرين «*him*»، و«*us*» في جملة «*He hit him*»، وعبارة «*with us*» على التوالي. أما القسم الثاني فيضم الضمائر الموصولة (الأسماء الموصولة) *Relative pronouns*، مثل «الذي»، و«التي»، و«الذين».... إلخ. ويضم القسم الثالث ضمائر الإشارة (أسماء الإشارة) *Demonstrative pronouns*، مثل «هذا» و«هذه»، و«هؤلاء». يوضح الشكل رقم (٩) التقسيم الفرعي للضمائر.



الشكل رقم (٩)

التقسيم الفرعي للضمائر

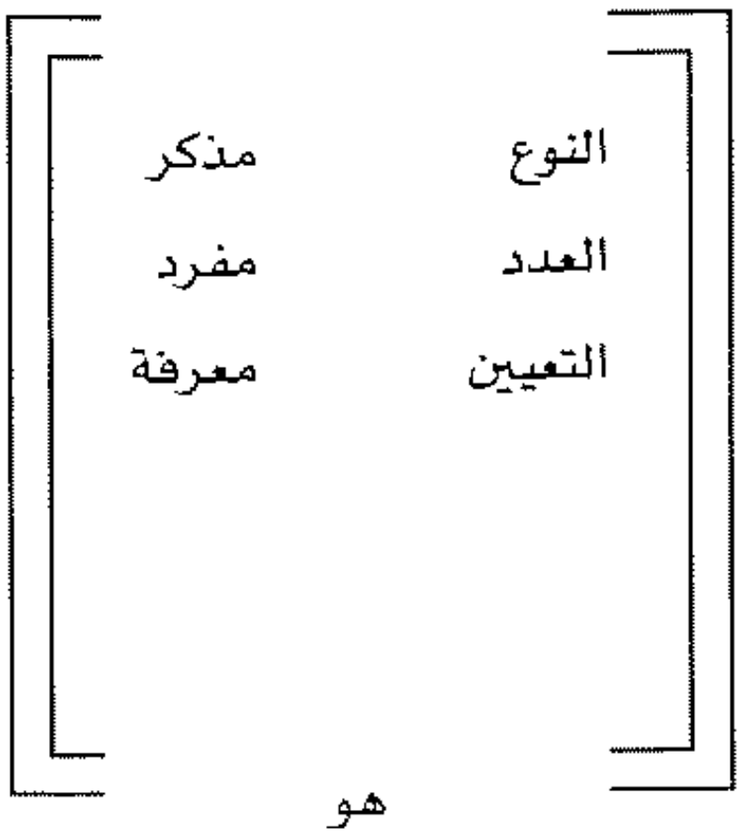
تختلف الضمائر باعتبارات: الشخص، والنوع، والعدد، على النحو الموضح بالجدول التالي:

الجدول رقم (٧)

حالات تصريف الضمير

الشخص		
المتكلم	المخاطب	الغائب
النوع		
مذكر	مؤنث	مذكر/مؤنث
العدد		
مفرد	مثنى	جمع

وعلى هذا يكون التمثيل المعجمي للمحتوى الصرفي للضمائر مشتملًا على المعلومات الصرفية الخاصة بالشخص، والنوع، والعدد، كما يوضح التمثيل المعجمي للمحتوى الصرفي للضمير «هو» بالشكل رقم (١٠).



الشكل رقم (١٠)

التمثيل المعجمي للمحتوى الصرفي للضمير «هو»

الأفعال المساعدة:

(١) كان وأخواتها:

تستخدم «كان» في زمن الماضي مع الأفعال الرئيسية في زمن المضارع للتعبير عن زمن الماضي المستمر، ويتصرف الفعلان تبعًا للشخص والنوع والعدد. يقابل هذا التركيب في اللغة الإنجليزية فعل الكينونة في زمن الماضي والفعل الرئيسي مضافًا إليه «*ing*» أو الفعل «*used to*».[١] كما في الأمثلة التالية:

كان يشاهد المباراة	*He was watching the match*
كانت تسير مع أختها	*She was walking with her sister*

وللتعبير عن عادة في الماضي:

كانوا يستيقظون يوميًا في السابعة	*They used to wake up at 7 daily*

كما تستخدم «كان» مع «قد» والفعل الرئيسي في زمن الماضي للتعبير عن صيغة التمام في الماضي التي يقابلها في الإنجليزية «*had + pp*"، في حالة البناء للمعلوم، و«*had +verb+ pp*» في حالة البناء للمجهول، كما في المثالين التاليين:

كان الطالب قد اجتاز الامتحان	*The student had passed the exam*
كان الاجتماع قد أُلغي	*The meeting had been cancelled*

(٢) أفعال الوجوب والاحتمالية:

تستخدم اللغة العربية مجموعة من الأفعال للتعبير عن الضرورة والإلزام والاحتمالية. تدخل هذه الأفعال -عادة- في تركيب مع «أن» المصدرية والفعل في زمن المضارع، أو ما يعرف في العربية بالمصدر المؤول. هذه الأفعال هي «يجب»، و«ينبغي»، و«يتعين»، و«يقتضي»، و«يحتاج» و«يستلزم»، و«يتطلب»، و«يستدعي»، و«يتوجب» للإلزام. تقابل هذه المجموعة في اللغة الإنجليزية الأفعال «*have to*»، و«*must*»، «*ought to*»، أو التعبير «*It is necessary*» في حين يستخدم الفعل «يمكن» للاحتمالية والاستطاعة ويقابله في الإنجليزية «*may*»، أو«*could*»، أو«*can*» أو التعبير«*It is possible*»على النحو الموضح بالأمثلة التالية:

(1) *Ryding, K. (2005). A Reference Grammar of Modern Standard Arabic. New York: Cambridge University Press. P. 446 - 447.*

يجب علينا أن نتعاون من أجل بلدنا *We have to cooperate for the sake of our country*

ينبغي علينا أن ندعم المجتهدين *We ought to support hard workers*

يمكن أن نحل المشكلة *We can solve the problem*

(٣) أفعال المقاربة والرجاء والشروع والاستمرارية:

تستخدم العربية مجموعة من الأفعال للدلالة على اقتراب وقوع حدث، أو رجاء حدوثه، أو الشروع في عمله. وتدخل هذه الأفعال في تركيب مع أفعال أخرى، أو مع «أن» المصدرية والفعل في زمن المضارع، مما قد يؤدي إلى اختلاف دلالة الفعل حال دخوله في هذا التركيب، كما في اختلاف دلالة الفعل «أخذ» في جملة «أخذ الكتاب» عنها في جملة «أخذ يجري».

كاد يسقط. *He was about to fall.*

أخذ يجري. *He started running.*

بدأ يذاكر دروسه *He started studying his lessons.*

كلمات الانفعال:

يقصد بكلمات الانفعال مجموعة الكلمات التي تعبر انفعال خالص(١). هذه الكلمات ترد عادة في سياقات معينة كالترحيب «مرحبًا»، و«أهلًا»، والإجابة «نعم»، و«لا»، و«كلا». كما تشتمل هذه المجموعة على مجموعة الكلمات التي يُطلَق عليها «أسماء الأفعال» مثل «صه»، و«أف»، وكذلك ما يطلق عليه «اسم الصوت»، مثل «كخ»، وكلمتي التعجب «نعم» و«بئس»(٢).

---

(1) Crystal, D. (2008). P. 249.

(٢) يطلق تمام حسان على هذه الكلمات اسم الخالفة، وقد تناولها بشيء من التفصيل في ضوء تصنيفه للكلام باعتباري المعنى والمبنى. راجع: حسان، تمام. (١٩٩٨). ص ١١٣ – ١١٨.

الحروف:

الحروف هي مجموعة من الكلمات التي لا تحمل معنى في ذاتها، وتكون خالية من التصريف، وتستخدم لأداء وظيفة معينة. تنقسم الحروف إجمالًا إلى ثلاثة أقسام: حروف مختصة بالاسم *Noun particles*، كحروف الجر «من، و«في»، و«عن»، و«إلى» ... إلخ، و«أل» التعريف، وكذلك حروف النداء، والاستثناء. وأخرى مختصة بالفعل *Verb particles*، كالسين التي للاستقبال و«سوف»، و«قد»، و«لقد»، وثالثة تدخل على الأسماء والأفعال كحروف العطف. تتفرع هذه الحروف بعد ذلك تبعًا لوظائفها النحوية أو الدلالية.

الرموز المستخدمة في التصنيف:

خصصت الدراسة مجموعة من الرموز لتمثيل الأنواع المعجمية المختلفة وكذلك الخصائص الصرفية والنحوية والدلالية داخل المعجم، على النحو التالي:

أولاً: رموز الأنواع المعجمية:

م	الرمز	التوصيف	أمثلة
١	+ *Noun*	الاسم	رجل – شجرة – سعادة – ذكاء – كتيب
٢	+ *Prop*	العلم	محمد – القاهرة – غزة
٣	+ *ActPart*	اسم الفاعل	قائم – ساجد – راكع
٤	+ *PassPart*	اسم المفعول	مسروق – معلوم – مضروب
٥	+ *Verbal*	المصدر	وصول – إدارة – ضرب
٦	+ *Adj*	صفة نعتية	حزين – جميل – قبيح
٧	+ *Comp*	صفة مقارنة	أفضل – أسعد – أنقى
٨	+ *Adv*	الظرف	أمس – اليوم – خلف – أمام
٩	+ *Rel*	الظرف الموصول	أنى – أيان – أينما – حيث
١٠	+ *Tim*	ظرف زمان	آنذاك – الآن – أمس
١١	+ *Loc*	ظرف مكان	أمام – خلف – تحت – فوق

١٢	+ *Verb*	الفعل	كتب – لعب – ضرب
١٣	+ *Aux*	الفعل المساعد	كان – كاد – ليس
١٤	+ *Num*	العدد	واحد – أربع – الثالث – السابع
١٥	+ *Card*	العدد الأصلي	واحد – أربع – خمسين – مليون
١٦	+ *Ord*	العدد الترتيبي	الثالث – السابع – العاشر
١٧	+ *Quant*	كلمات الكم	كل – بعض – بضعة– شَطْر
١٨	+ *Pron*	الضمير	أنا – هو – إياك – هذا – الذي
١٩	+ *Sub*	ضمير شخصي (فاعل)	أنا – هو – أنت – نحن – هما
٢٠	+ *Obj*	ضمير شخصي (مفعول)	إياي – إياك – إيانا – إياهم
٢١	+ *Pos*	ضمير ملكية	ي – نا – هم – كما
٢٢	+ *Dem*	ضمير (اسم إشارة)	هذا – هذه – هذان – هؤلاء
٢٣	+ *Rel*	ضمير (اسم) موصول	الذي – التي – الذين – اللائي
٢٤	+ *Art*	الأداة	إن – إلا – كي – يا
٢٥	+ *Interj*	كلمات الانفعال	صه – أف – أهلا – نعم
٢٦	+ *Prep*	الجر	من – إلى – على – من – في
٢٧	+ *Conj*	العطف	و – ف – ثم – أو – أم
٢٨	+ *SubConj*	ربط الجمل	إنْ – لو – لولا
٢٩	+ *Voc*	النداء	الهمزة، يا، أيْ، أية
٣٠	+ *AccArt*	النصب	أنْ – لن – حتى – كي
٣١	+ *Exp*	الاستثناء	إلا – غير – سوى – حاشا – عدا
٣٢	+ *Neg*	النفي	لا، لم، لن، لمّا
٣٣	+ *Pseudo_Verb*	المشبهة بالفعل	إنَّ وأخواتها، ولا النافية للجنس
٣٤	+ *Equal*	التسوية	الهمزة
٣٥	+ *Conf*	التوكيد	اللام
٣٦	+ *Interp*	التفسير	أيْ
٣٧	+ *Rest*	الحصر	إنما، أنما
٣٩	+ *Interrog*	الاستفهام	هل، الهمزة

٤٠	+ *Focus*	التركيز	أمَّا
٤١	+ *Proh*	النهي	لا
٤٢	+ *Oath*	القسم	الواو، الباء، التاء
٤٣	+ *RC*	جواب الشرط	الفاء، اللام
٤٤	+ *JussArt*	الجزم	لم، لمّا، لا
٤٥	+ *Inf*	مصدري	أنْ
٤٦	+ *Det*	محدد	أل التعريف
٤٧	+ *Siml*	تشبيه	كأن
٤٨	+ *Hope*	التمني	ليت
٤٩	+ *Beg*	الرجاء	لعل
٥٠	+ *Amend*	الاستدراك	لكن
٥١	+ *Cond*	الشرط	إن
٥٢	+ *Exclm*	التعجب	ما
٥٣	+ *Caus*	السببية	الفاء
٥٤	+ *Resu*	الاستئناف	الواو
٥٥	+ *Doubt*	الشك	قد
٥٦	+ *Cert*	التوكيد	قد – لقد
٥٧	+ *Ret*	الإضراب	بل
٥٨	+ *Fut*	الاستقبال	السين
٥٩	+ *Exh*	التحضيض	ألا – لولا
٦٠	+ *GenArt*	أداة جر	سوى

ثانيًا: رموز الخصائص الصرفية النحوية:

الرمز	التوصيف	أمثلة
(١) الموقع الإعرابي للاسم		
+ Nom	الرفع	كتابٌ – رجلٌ – محمدٌ
+ Acc	النصب	كتابًا – رجلاً – محمدًا
+ Gen	الجر	كتابٍ – رجلٍ – محمد
(٢) الحالة الإعرابية للفعل		
+ Ind	الرفع	يكتبون – يلعبون – يجري
+ Subj	النصب	يكتبوا – يلعبوا – ينتهي
+ Juss	الجزم	يكتبوا – يلعبوا – يجر
(٣) الزمن		
+ Perf	الماضي	كتب – لعب – نام
+ Impf	المضارع	يكتب – يلعب – ينام
+ Impv	الأمر	اكتب – العب – نم
(٤) العدد		
+ Sg	المفرد	طالب – معلم – فتاة
+ Dual	المثنى	طالبان – معلمان – فتاتان
+ Pl	الجمع	طلاب – معلمون – فتيات
+ Mass	غير معدود	لبن – سكر – ماء
(٥) الجنس		
+ Masc	مذكر	محمد – كتاب – هو
+ Fem	مؤنث	فاطمة – ساعة – هي
+ FM	مذكر / مؤنث	عجوز
(٦) بناء الفعل		
+ Act	المعلوم	كتب – يكتب – ضرب – يضرب
+ Pass	المجهول	كُتِب – يُكتَب – ضُرِب – يُضرَب

(٧) الشخص		
+ *P1*	المتكلم	أنا – نحن – كتبت – لعبنا
+ *P2*	الغائب	هو – هي – كتب – يكتب – لعبت
+ *P3*	المخاطب	أنت – انتِ – تكتب – تلعبين
(٨) التعريف		
+ *Def*	المعرفة	الكتاب – القلم – الفتاة
+ *Indef*	النكرة	كتاب – قلم – فتاة
(٩) التعدي		
+ *Trans*	متعد	أعطى – منع – كلّم
+ *Intrans*	لازم	مات – نام – ذهب

ثالثًا: رموز الخصائص الدلالية:

+ *Abst*	مجرد	سعادة – حب – تعاون
+ *Con*	ذات	شجرة – كتاب – باب
+ *Hum*	بشري	طالب – معلم – فتاة
+ *Inhum*	غير بشري	شارع – جامعة – منزل
+ *Geo*	كيان جغرافي	القاهرة – مكة – القدس
+ *Org*	منظمة	جوجل – يونيسكو – أوراكل
+ *Doc*	وثيقة	بلفور – كامب ديفيد
+ *Tool*	أدوات	جافا – نوكيا – ويندوز
+ *Scient*	علمي	سرطان – انترفيرون
+ *Temp*	زمني	الأحد – مايو – محرم
+ *Dif*	مختلف	الله

المدونة المستخدمة:

اعتمدت الدراسة في استقاء المداخل المعجمية على المدونة اللغوية[1] التي طورتها لطيفة السُّليطي في أطروحتها لنيل درجة الماجستير من جامعة ليدز، وهي مدونة تحتوي على حوالي مليون كلمة. ويرجع اعتماد الدراسة على هذه المدونة تحديدًا إلى كونها مصصمة -بشكل أساسي- لأغراض ثلاثة، أحدها «هندسة اللغة»، كما أنها تلبي حاجة المعجم المقترح من النصوص التي تعكس الاستخدام المعاصر للغة العربية؛ وهي النصوص التي اعتمدت الباحثة في تحديدها على استبيان مساعد؛ حتى لا يتم اختيار نصوص المدونة عشوائيًا. أما السبب الثالث فيرجع إلى كونها متاحة[2] مجانًا على شبكة الإنترنت، وذلك بعد أن حصلت الباحثة على الإذن القانوني من المصادر،[3] بخلاف مدونات أخرى إما أنها غير متاحة مجانًا، أو أنها جُمِعت لغرض بحثي محدد، وبذلك لا تعد مصدرًا عامًّا. فإذا أضفنا إلى ذلك التصميمَ الآليَّ، تكون بذلك قد استوفت أهم خصائص المدونة المطلوبة، وهي:

- التصميم الآلي، أو إمكانية القراءة الآلية.

---

(١) يقصد بالمدونة *Corpus* - من وجهة النظر اللغوية- تجميع لمادة نصية أو كلامية وفق قواعد أو معايير محددة. وللمدونات اللغوية أهمية كبيرة في مجالات البحث اللغوي - النظري والتطبيقي- كدراسات القواعد، والمعاجم، والتنوع اللغوي، واللغويات التاريخية، والتحليل التقابلي، ونظرية الترجمة، والمعالجة الآلية للغات الطبيعية. وقد أسهمت التطورات التقنية الهائلة في تيسير عمليات التخزين الحاسوبي، والدخول إلى كميات كبيرة من المعلومات؛ فزادت بصورة واضحة فكرة بناء المدونات، واستخدامها في المجالات البحثية المختلفة.

(٢) قام الباحث بتحميل المدونة عبر الرابط التالي:

http://www.comp.leeds.ac.uk/eric/latifa/research.htm

(3) Al-Sulaiti, L. (2004). The Introduction of: Designing and Developing a Corpus of Contemporary Arabic. A M.A. thesis, University of Leeds, School of Computing.

❖ تمثيل المجال محل الدراسة: ويتم ذلك من خلال تحقيق درجة دقة مقبولة، وصورة نسبية للقطاع اللغوي كله.

❖ الاعتماد على عينة متوازنة: ويتحقق ذلك باشتمال المدونة على عينات من أنواع متنوعة (شعر – نثر – أخبار – رسائل إلكترونية... إلخ).

تجدر الإشارة إلى أن الاستبيان الذي أعدته الباحثة قد اشتمل على سؤال يتعلق بتحديد التطبيقات المستقبلية الأساسية للمدونة، وكانت التطبيقات الأساسية التي اقترحها مهندسو اللغة هي: بناء نظم الترجمة الآلية وتطويرها، وبناء نظم استخلاص المعلومات، والمدققات النحوية، وتعرُّف الكلام، وإنتاج الكلام، وتحويل الكلام إلى نص، والعكس.(١)

تشتمل المدونة -تبعًا لتوصيف الباحثة- على ثمانمائة واثنتين وأربعين ألفًا وستمائة وأربع وثمانين (٨٤٢,٦٨٤) كلمة في أربعمائة وستة عشر (٤١٦) ملفًا، جُمعت من خمسة مصادر، هي: الجرائد، والمجلات، والإذاعة، والتلفاز، وصفحات الإنترنت، توزعت عليها أنواع النصوص بنسب مختلفة، على النحو الموضح بالجدول رقم (٨).

---

(1) Ibid. P.30.

الجدول رقم (٨)

أنواع النصوص التي اعتمدت عليها المدونة المستخدمة، وعدد ملفاتها وكلماتها[١]

م	نوع النص	عدد النصوص	عدد الكلمات	ملاحظات
		مكتوب		
١	قصص قصيرة	٣١	٤٥,٤٦٠	
٢	تعليم	١٠	٢٥,٥٧٤	
٣	دين	١٩	١١١,١١٩	
٤	شعر	٥	١,١٤٧	
٥	ترفيه	٢	٤,٠١٤	
٦	سيرة ذاتية	٧٣	١٥٣,٤٥٩	
٧	علم اجتماع	٣٠	٨٥,٦٨٨	
٨	سياحة/سفر	٦١	٤٦,٠٩٣	
٩	وصفات	٩	٤,٩٧٣	
١٠	وثائق علمية	٤٥	١٠٤,٧٩٥	
١١	رياضة	٣	٨,٢٩٠	
١٢	اقتصاد	٢٩	٦٧,٤٧٨	
١٣	قصص أطفال	٢٧	٢١,٩٥٨	
١٤	صحة وطب	٣٢	٤٠,٤٨٠	
١٥	وثائق حاسوب وإنترنت	٢	١٢,٢٩٧	
١٦	مقابلات شخصية	٢٤	٥٨,٤٠٨	

---

(١) أخذت الباحثة نصوص المدونة من عدد من المصادر -بعد الحصول على إذن قانوني- ومنها: مجلة العربي، ومجلة أفق، والحرية، والقصة العربية، وأخبار اليوم، وعالم الاقتصاد، ومجلة المعرفة، وموقع إسلام أون لاين، وغير ذلك. للمزيد، المرجع السابق. ص ٥٤ – ٥٩.

١٧	سياسة	٩	٤٦,٢٩١	
		منطوق		
١٨	تعليم	٢	١,٢٤٠	
١٩	رياضة	٣	١,٧٣٦	مستوى تعليمي
٢٠	ترفيه	١	١,٣٧٧	مستوى عامي
٢١	سياسة	١	١,٢٥٢	

الكلمات الأكثر تكرارًا في المدونة:

بعد الانتهاء من تحميل المدونة، اعتمدت الدراسة على مُفهرَس يطلق عليه «*tlCorpus*[(١)]» طوره ديفيد جوفي *David Joffe*، ومالكولم ماكليود *Malcolm Macleod* لإعداد قائمة بمفردات المدونة، وعدد مرات تكرارها؛ بهدف تطبيق المنهجية المقترحة على مجموعة الكلمات الأكثر تكرارًا، مع مراعاة اشتمالها على كل الأنواع المعجمية. ومن اللافت -أولًا- أن المفهرس المستخدم قد أظهر أن عدد كلمات المدونة أقل بكثير من العدد الذي ذكرته الباحثة، فأظهر أن عدد الأنواع سبعون ألفًا وتسعمائة وثلاثة (٧٠,٩٠٣) أنواع، وأن عدد الوحدات اللغوية المتمايزة ثلاثمائة وثمانية وعشرون ألفًا وتسعمائة وثماني وخمسون (٣٢٨,٩٥٨) وحدة. ومن اللافت كذلك أن المفهرس لم يتعامل مع الكلمة حال استقلالها عن السوابق واللواحق واللواصق، وحال اتصالها بها بوصفهما شيئًا واحدًا، ومن ثم عَد «في» كلمة، و«وفي» كلمة أخرى، و«فيها» كلمة ثالثة، وهكذا، ورتب كل واحدة منها بحسب تكرارها في المدونة. وقد رد الباحث هذه الكلمات إلى جذعها عند إدخالها في المعجم. ويحتوي الجدول رقم (٩) على قائمة بأكثر مائة كلمة مكررة في المدونة.

---

(١) قام الباحث بتحميل نسخة تجريبية عبر الرابط التالي:

http://tshwanedje.com/

الجدول رقم (٩)

قائمة بأكثر مائة كلمة تكرارًا في المدونة

م	الكلمة	مرات التكرار	النسبة المئوية	م	الكلمة	مرات التكرار	النسبة المئوية
١	في	١٠٤٥٢	٣,١٨٪	٢	مِن	٨٣٤٣	٢,٥٤٪
٣	على	٤٢٦١	١,٣٠٪	٤	أن	٤٢٣٢	١,٢٩٪
٥	إلى	٣٥٨٧	١,٠٩٪	٦	التي	٢٠٦٧	٠,٦٣٪
٧	عن	١٨٧٩	٠,٥٧٪	٨	ما	١٨٤٦	٠,٥٦٪
٩	لا	١٧٨٢	٠,٥٤٪	١٠	أو	١٦٨٣	٠,٥١٪
١١	هذا	١٦٠٦	٠,٤٩٪	١٢	الذي	١٤٩٣	٠,٤٥٪
١٣	هذه	١٤١٧	٠,٤٣٪	١٤	كان	١٤٠٢	٠,٤٣٪
١٥	مع	١٠٢٥	٠,٣١٪	١٦	كل	٩٨٩	٠,٣٠٪
١٧	لم	٩٤٧	٠,٢٩٪	١٨	هو	٩١٤	٠,٢٨٪
١٩	ذلك	٨٨٨	٠,٢٧٪	٢٠	بين	٨٨٠	٠,٢٧٪
٢١	بعد	٨٢٢	٠,٢٥٪	٢٢	كما	٧٧٠	٠,٢٣٪
٢٣	حتى	٧١٨	٠,٢٢٪	٢٤	كانت	٦٨٣	٠,٢١٪
٢٥	قد	٦٥٤	٠,٢٠٪	٢٦	إن	٦٥٢	٠,٢٠٪
٢٧	وقد	٦٤٦	٠,٢٠٪	٢٨	عام	٥٩٤	٠,١٨٪
٢٩	هي	٥٨٠	٠,١٨٪	٣٠	وفي	٥٦١	٠,١٧٪
٣١	ثم	٥٣٩	٠,١٦٪	٣٢	العربية	٥١٨	٠,١٦٪
٣٣	وهو	٥١٠	٠,١٦٪	٣٢	له	٤٧٠	٠,١٤٪
٣٣	حيث	٤٦٦	٠,١٤٪	٣٤	ولا	٤٦١	٠,١٤٪
٣٥	أنه	٤٥١	٠,١٤٪	٣٦	الله	٤٤٩	٠,١٤٪
٣٧	بعض	٤٤٥	٠,١٤٪	٣٨	إلا	٤٣٦	٠,١٣٪
٣٩	وكان	٤٣٦	٠,١٣٪	٤٠	غير	٤٣٠	٠,١٣٪
٤١	العالم	٤٢٧	٠,١٣٪	٤٢	هناك	٤١٠	٠,١٢٪

٤٣	أي	٤٠٥	٠,١٢٪	٤٤	فيها	٤٠٣	٠,١٢٪
٤٥	ومن	٤٠٢	٠,١٢٪	٤٦	أكثر	٣٩٦	٠,١٢٪
٤٧	خلال	٣٨٤	٠,١٢٪	٤٨	عليه	٣٨٤	٠,١٢٪
٤٩	العربي	٣٧٨	٠,١١٪	٥٠	قبل	٣٦٧	٠,١١٪
٥١	بل	٣٥٨	٠,١١٪	٥٢	ولكن	٣٥٧	٠,١١٪
٥٣	فيه	٣٤٨	٠,١١٪	٥٤	تلك	٣٤٦	٠,١١٪
٥٥	عند	٣٤٥	٠,١٠٪	٥٦	يمكن	٣٤١	٠,١٠٪
٥٧	الذين	٣٣٨	٠,١٠٪	٥٨	به	٣٢٦	٠,١٠٪
٥٩	مثل	٣١٤	٠,١٠٪	٦٠	أنا	٣٠٩	٠,٩٪
٦١	أحد	٣٠٨	٠,٠٩٪	٦٢	وهي	٣٠٧	٠,٩٪
٦٣	فقد	٣٠٢	٠,٠٩٪	٦٣	بها	٣٠١	٠,٩٪
٦٤	منها	٢٩٩	٠,٠٩٪	٦٥	يكون	٢٩٧	٠,٩٪
٦٦	عندما	٢٩٦	٠,٠٩٪	٦٧	الآن	٢٩٣	٠,٩٪
٦٨	الإنسان	٢٩٢	٠,٠٩٪	٦٨	لي	٢٨٧	٠,٩٪
٦٩	دون	٢٨٤	٠,٠٩٪	٧٠	لها	٢٨٤	٠,٩٪
٧١	منذ	٢٨٤	٠,٠٩٪	٧٢	أما	٢٨٣	٠,٩٪
٧٣	إذا	٢٨١	٠,٠٩٪	٧٤	أخرى	٢٧٠	٠,٨٪
٧٥	سنة	٢٧٠	٠,٠٨٪	٧٦	يا	٢٦١	٠,٨٪
٧٧	كنت	٢٦٠	٠,٠٨٪	٧٨	العمل	٢٥٦	٠,٨٪
٧٩	الناس	٢٥٢	٠,٠٨٪	٨٠	الإسلام	٢٥٠	٠,٨٪
٨١	وما	٢٥٠	٠,٠٨٪	٨٢	حين	٢٤٥	٠,٧٪
٨٣	آخر	٢٤٠	٠,٠٧٪	٨٤	لقد	٢٣٩	٠,٧٪
٨٥	قال	٢٣٦	٠,٠٧٪	٨٦	فإن	٢٣٤	٠,٧٪
٨٧	ولم	٢٣٣	٠,٠٧٪	٨٨	لكن	٢٣١	٠,٧٪
٨٩	التعليم	٢٩٩	٠,٠٧٪	٩٠	ليس	٢٢٧	٠,٧٪
٩١	منه	٢١٨	٠,٠٧٪	٩٢	هنا	٢١٥	٠,٧٪
٩٣	الحياة	٢١٣	٠,٠٦٪	٩٤	مما	٢١٢	٠,٠٦٪

٩٥	يجب	٢٠٩	٠,٠٦٪	٩٦	محمد	٢٠٧	٠٦٪
٩٧	نفسه	٢٠٧	٠,٠٦٪	٩٨	عليها	٢٠٤	٠٦٪
٩٩	الشعر	٢٠١	٠,٠٦٪	١٠٠	إليه	١٩٩	٠٦٪

وصف المعجم:

طور الباحث معجمًا لتطبيقات المعالجة الآلية للعربية المعاصرة، وقد زود المعجم بمحلل صرفي ثنائي الاتجاه *Bidirectional Morphological Analyze*، باستخدام تقنية الحالة المحدودة *Finite State Technology*؛ لتوليد الصيغ المتصرفة آليًا. ويمكن وصف المعجم على النحو التالي:

- ❖ معجم آلي عام ثنائي اللغة، قوامه حوالي (٢٠٠٠) مدخل.
- ❖ مداخل المعجم مستخرجة من مدونة عامة، وهي الكلمات الأكثر تكرارًا في المدونة.
- ❖ موجه – أساسا– لتطبيقات المعالجة الآلية للغة العربية المعاصرة.
- ❖ يعتمد في بنائه على الجذع دون الجذر والوزن.
- ❖ يشتمل المدخل على الصيغ غير المنتظمة انتظامًا تامًا.
- ❖ يحتوي على محلل صرفي آلي ثنائي الاتجاه لتوليد الصيغ المنتظمة انتظامًا تامًا.
- ❖ يعالج جوانب القصور في محتوى المعاجم الآلية الحالية.
- ❖ قابل للتوسيع.

اتخذت الدراسة من المعجم الحاسوبي العربي المصمم لأغراض تكنولوجيا اللغة موضوعًا لها؛ بهدف بناء معجم آلي يلبي متطلبات نظم معالجة اللغة آليًا، ويتجاوز جوانب القصور في المعاجم المصممة لأغراض الاستخدام البشري، وقد زود بمحلل صرفي لتوليد الصيغ المنتظمة آليًا.

انطلقت الدراسة من مجموعة من الأفكار اللسانية، واللسانية الحاسوبية، توزعت جوانبها على جوانب العمل في بناء المعجم المقترح؛ فعلى مستوى التأسيس النظري اقترحت الدراسة اعتماد الجذع اللغوي أساسًا لبناء المعجم،

كما اقترحت تصنيفًا للأنواع المعجمية العربية (أقسام الكلام) راعت فيه بعض الخصائص التي تجاهلتها التصنيفات الحاسوبية الحالية للكلمة العربية، كالتمييز على المستوى الدلالي بين أسماء ذات وأسماء معاني، وكذلك الأسماء التي تشير إنسان *Human* مثل «رجل»، و«فتاة»، و«سفير»، و«مدير»، و«أستاذ»، في مقابل الأسماء التي لا تشير إلى إنسان *Inhuman* مثل «مسجد»، و«كتاب»، و«راية»... إلخ.

أما على مستوى التطبيقات المقترحة قدمت الدراسة تصميمًا لمعجم حاسوبي أُطلِق عليه (مريم كوملكس *MariamComLex*)، مزود بمحلل صرفي ثنائي الاتجاه يتجاوز جوانب قصور عديدة شابت نظم التحليل الصرفي العربية الحالية كالاكتفاء بالاعتماد على المعاجم التقليدية، وعدم الاعتماد على مدونة لغوية تمثل اللغة العربية المعاصرة في استخراج المواد المعجمية، والاعتماد على فكرة الجذع وليس على فكرة الجذور والأوزان، بالإضافة إلى جوانب قصور تتعلق بالتصنيف النحوي للأنواع المعجمية «أنواع الكلام»، وأخرى تتعلق بالمعلومات اللغوية المخزنة في المعجم.

ويمكن - بإيجاز- بلورة نتائج الدراسة فيما يلي:

- برهنت الدراسة على عدم دقة الفرضية القائلة بأن «المشتقات المختلفة للجذر الواحد تشترك -ولو جزئيًا- في معنى أساسي يحمله الجذر» وهي فرضية اتخذت بموجبها معظم المعاجم العربية - منذ بداية تأليفها - جذر الكلمة أساسًا تورد تحته كافة أنواع المشتقات، وأكدت الدراسة أن «اشتراك بعض الكلمات في حروف المادة له دلالته على الاتصال الشكلي بين هذه الكلمات، دون أن يعني - بالضرورة - اتصالًا بينها على مستوى الدلالة.
- خلصت الدراسة إلى أن اعتماد الجذع أساسًا لبناء المعجم الحاسوبي العربي يجعله أكثر فاعلية وسهولة في التطوير والتوسيع؛ إذ يساعد على التخلص من عملية التجذيع، وكذلك من الصيغ المهملة، ويمكن ربطه بالمعلومات اللغوية على مستوى الصرف والنحو والدلالة.
- قدمت الدراسة تصنيفًا جديدًا للأنواع المعجمية العربية راعت فيها عددًا من الخصائص التي أغفلتها التصنيفات الحاسوبية الحالية للكلمة العربية كما

أبرزت أهمية هذه الخصائص عند استخدام المعجم في التطبيقات الكبرى للمعالجة الآلية.

- قَدّمَتْ اقتراحًا وتصميمًا لمعجم حاسوبي مصمم لتطبيقات المعالجة الآلية للغة العربية، هو معجم «مريم كوملكس»، وهو مزود بمحلل صرفي ثنائي الاتجاه، اعتمادًا تقنية الحالة المحدودة.

قائمة المراجع

المراجع العربية:

- أبوالمكارم، علي. (١٩٨٠). تصريف الأسماء. الطبعة الأولى. المكتبة النحوية.
- أبوسكين، عبدالحميد محمد. (١٩٨١). المعاجم العربية: مدارسها ومناهجها. الطبعة الثانية. القاهرة: الفاروق الحديثة للطباعة والنشر.
- إقبال، أحمد الشرقاوي. (١٩٩٣). معجم المعاجم. الطبعة الثانية. بيروت: دار الغرب الإسلامي.

الباتلي، أحمد بن عبدالله. (١٩٩٢). المعاجم اللغوية وطرق ترتيبها. الرياض: دار الراية. الطبعة الأولى.

- الخطيب، عدنان. (١٩٩٤). المعجم العربي بين الماضي والحاضر. الطبعة الثانية. لبنان: مكتبة لبنان ناشرون.
- حجازي، محمود فهمي. (د.ت). البحث اللغوي. القاهرة: دار غريب للطباعة والنشر.
- حسان، تمام (١٩٩٨). اللغة العربية: معناها ومبناها. الطبعة الثالثة. القاهرة: عالم الكتب.
- رسلان، صلاح بسيوني. (د. ت.). كونفوشيوس رائد الفكر الإنساني. نسخة إلكترونية.
- فايد، وفاء كامل. (٢٠٠٧). واقع المعجمية الحاسوبية العربية. ندوة (العربية في عصر العولمة)، المجلس الأعلى للثقافة، ومجمع اللغة العربية بالقاهرة (في إطار الاحتفال باليوبيل الماسي لمجمع اللغة العربية)- القاهرة مارس ٢٠٠٧.
- مجمع اللغة العربية بالقاهرة. (٢٠٠٤). المعجم الوسيط. الطبعة الرابعة. القاهرة: مكتبة الشروق الدولية.

مهديوي، عمر. (٢٠٠٧). المعجم العربي في ضوء اللسانيات الحاسوبية. الندوة الدولية الأولى عن الحاسب واللغة العربية. مدينة الملك عبدالعزيز للعلوم والتقنية. الرياض، الفترة من ١٠ – ١٢ نوفمبر ٢٠٠٧.

- موسى، علي حلمي. (١٩٧٨). مقدمة: دراسة إحصائية لجذور معجم الصحاح باستخدام الكمبيوتر. الهيئة المصرية العامة للكتاب.
- نصار، حسين. (١٩٨٨). المعجم العربي: نشأته وتطوره. الطبعة الرابعة. الجزء الأول. القاهرة: دار مصر للطباعة.

المراجع الأجنبية:

- *Al-Sulaiti, L. (2004). The Introduction of: Designing and Developing a Corpus of Contemporary Arabic. A M.A. thesis, University of Leeds, School of Computing.*
- *Brisco, T. (1991). Lexical issues in Natural Language Processing. In Klein, E. & Veltman, F. (1991). Natural Language and Speech. Springer, Verlag.*
- *Brockington, J. (2006). Sanskrit. In Brown, K. & Ogilvie, S. (ed). (2009). Concise Encyclopedia of Languages of the World. First edition. Elsevier Ltd; Oxford, UK.*
- *Crystal, D. (2008). A Dictionary of Linguistics and Phonetics. Sixth edition. Blackwell Publishing Ltd.*
- *Dichy, J. & Farghaly, A. (2007). Grammar-Lexis Relation in The Computational Morphology of Arabic. In Soudi, A., Van Den Bosch, A. & Neumann, G. (eds.) (2007). Arabic Computational Morphology: Knowledge-based and Empirical Methods. Springer.*
- _____, *(2003) Roots & Patterns vs. Stems plus Grammar-Lexis Specifications: on what basis should a multilingual lexical database centred on Arabic be built? MT Summit IX - Workshop: Machine Translation for Semitic Languages. New Orland, USA.*
- *Dichy, J. & Hassoun, M. (eds). (1989). Simulation de modèles linguistiques et Enseignement Assiisté par Ordinateur de l arabe - Travaux SAMIA I.*

Paris, Conseil International de la langue Francaise.

- *Farghaly, A. (1987). Three Level Morphology. Paper presented at the Arabic Morphology Workshop, Linguistic Summer Institute, Stanford, CA.*
- *Farghaly, A. & Senellart, J. (2003). Intuitive Coding of the Arabic Lexicon. In the Proceedings of MT Summit IX, Workshop on Machine Translation for Semitic Languages: Issues and Approaches, New Orleans, September 2003.*
- *Habash, N. (2010). Introduction to Arabic Language Processing. Morgan & Claypool Publishers Series.*
- *Habash, N & Rambow, O. (2006). MAGEAD: A Morphological Analyzer and Generator for Arabic Dialects. In Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meetings of the Association for Computational Linguistics. pp. 681 - 688. Association for Computational Linguistics, Australia: Sydne.*
- *Haspelmath, M. (2001). Word Class and Parts of Speech. In the International Encyclopedia of Social & Behavioral Sciences. Elsevier Science Ltd.*
- *Jurafsky, D. & Martin, J. (2006). Speech and Language Processing: An Introduction to Language Processing.*
- *Megerdoomian, K. (2003). In Farghaly, A. (ed). Handbook for Language Engineers. Center for the Study of language and Information.*
- *Paik, W., Liddy, E., Yu, E. & McKenna, M. (1993b). Categorizing and Standardizing Proper Nouns for Efficient Information Retrieval. In B. Boguraev & J. Pustejovsky (Eds.), Proceedings of the workshop on acquisition of lexical knowledge from text, Ohio, USA (pp.154-160). Cambridge, MA: MIT Press.*
- *Ryding, K. (2005). A Reference Grammar of Modern Standard Arabic.*

New York: Cambridge University Press.

- *Sinclair, J. (ed.) 1987. Looking Up: An Account of the COBUILD Project in Lexical Computing. London: Collins.*
- *SmrZ, O. (2007). Functional Arabic Morphology: Formal System and Implementation. PhD Thesis, Charles University in Prague, Prague, Czech Republic.*
- *Sterkenburg, P. The Dictionary: Definition and History. In Sterkenburg, P. (2003). A Practical Guide to Lexicography . Terminology and Lexicography Research and Practice. ISSN 7067; V. 6. John Benjamin B.V.*
- *Wilson, R. & Frank C. (eds.). (1999). Computational Lexicons. In The MIT Encyclopedia of the Cognitive Sciences. The MIT Press; Cambridge; Massachusetts; London; England.*
- *Yao, X. (2000). An Introduction to Confucianism. Cambridge: Cambridge University Press.*