معالجة المحتوى المعجمي الدلالي في المعجم العربي الحاسوبي
مقاربة لغوية حاسوبية

[*1]أ.د. وفاء كامل، [**2]أ.د. محسن رشوان، [*]د. عبد العاطي هواري
[*]كلية الآداب، جامعة القاهرة
[**]كلية الهندسة، جامعة القاهرة
[1]wafkamel@yahoo.com
[2]mrashwan@rdi-eg.com

ملخص الورقة:

موضوع الورقة هو **مقترح لمنهجية معالجة المحتوى المعجمي الدلالي للمعجم العربي الحاسوبي** (في صورة تعريفات)؛ إذ يعد الجانب الدلالي أهم جوانب صناعة مصدر معجمي، لعدة أسباب هي:

- كونه الغايةَ الأساسية من أي مصدر معجمي، يستهدف المستعمل البشري، أو المعالجة الحاسوبية.
- تعَقّد جوانب المحتوى المعجمي واتساعها لتشمل جانبا من المحتوى الصرفي الدلالي، والجانب التركيبي، إضافة إلى الجانب الدلالي نفسه.
- أن المحتوى المعجمي الدلالي في المعجم هو مصدر المعالجات الحاسوبية، التي تهدف إلى الفهم الآلي للغة، ومعالجة جوانبها الدلالية.

وقد اقتُرِحت تصورات متعددة لصناعة مصادر معجمية عربية حاسوبية[1]، غير أن مقارباتها الجوانبَ المعجمية الدلالية جاءت تقليدية وغير معمقة، بل إنها تكاد تكون نَسْخا لمقاربات المعجم العربي الأساسي والمعجم الوسيط[2]، فضلا عن المعاجم العربية التراثية.

واللسانيون الحاسوبيون يقعون دائما في مشكلات عند تعاملهم الحاسوبي مع المصادر المعجمية العربية. كما أن المصادر الحاسوبية منها تقليديةٌ في معالجتها لهذا الجانب أيضا؛ لذا كان لزاما- عند إرادة البدء في صناعة مصدر معجمي عربي- البدء ببناء منهجية لمعالجة المحتوى المعجمي الدلالي، الذي يعد أهم مشكلات المعجم العربي عموما؛ إذ إن القضايا الدلالية في هذا المعجم قد عولجت دون منهج واضح.

وينبني **التصور النظري للمقترح** على عدة جوانب هي:

١. المنطلقات النظرية : الصرفية والتركيبية والدلالية.
٢. التحليل : مدونته، وإجراءاته، ومخرجاته.
٣. التمثيل : لغته، وأبنيته.

الكلمات المفتاحية:

المعجم العربي الحاسوبي Arabic Computational Lexicon - الصناعة المعجمية Lexicography– المصادر المعجمية العربية الحاسوبية Lexical Resources Arabic Computational - لغة التعريف المعجمي Lexical Definition Language- المحتوى المعجمي الدلالي العربي Lexical Semantic Arabic Content- المدونة المعجمية Lexical Corpus- الدلالة المعجمية الحاسوبية Computational Lexical Semantics- اللسانيات المعجمية Lexical Linguistics.

مدخل:

تُعَد المصادر المعجمية Lexical Resources ، على تنوعها، البنيةَ التحتيةَ Infrastructure لأنظمة معالجة اللغات الطبيعية Natural Language Processing واللسانيات الحاسوبية Computational Linguistics[3]؛ لأنها تقدم لها المعارف اللغوية التي تمثل لبنات الجملة والنص اللغوي. وقد "أصبح للمعجم، في مجال اللسانيات الحاسوبية، موقع محوري بل أصبح المعجم مفتاح الدلالة، ومحور التوليد اللغوي. وقد اقتضاه هذا الموقع أن يقنن منهج الشرح المعجمي على نحو يتجاوز ما كان متعارفا من أمر المعجمات التقليدية"[4]. فالمعجم يمثل عنق الزجاجة: من وجهة نظر جل العاملين بحوسبة اللغة.

وبالرغم من تنوع أشكال المصادر المعجمية، واختلافها في الأطر النظرية الموجهة لها، إلا أنها تتفق جميعها في الاهتمام بجانب المحتوى المعجمي الدلالي، مع اختلاف في طريقة تمثيله. إذ يقع المحتوى المعجمي الدلالي في قلب اهتمام معالجة اللغات الطبيعية باعتباره غايةَ أيِّ تواصلٍ لغوي، كما أن غاية أي مشروع معجمي، مهما كانت أسسه النظرية وغاياته العملية، هي تمثيل هذا المحتوى.

مبررات المقترح:

- وجود مشكلات في معالجة الجوانب المعجمية الدلالية في المعجم العربي التقليدي والحاسوبي، مردها غياب التصور النظري المتكامل للإشكالية، الذي يكون أساسا للتطبيق المعجمي[5] .
- أشارت معظم دراسات المعجم العربي إلى نقائص معالجة المحتوى المعجمي، وخلصت إلى ما يشبه التوجيهات لكل من يَرْغب في تأليف معجم. و من التوجيهات ما يخص التعريف[6]. على أنه لا يمكن الاكتفاء بالنقد دون التصدي لإنشاء معجم موسع، مبني على تحليل معجمي دلالي معمق، وتمثيل مصوغ بلغة تعريف منضبطة، يصلح أساسا للبحث والتوظيف في معالجة اللغات الطبيعية.

---

(١) تتبنى منظمة الأليكسو بالتعاون مع مدينة الملك عبد العزيز للعلوم والتقنية بالسعودية ومجمع اللغة العربية بدمشق مشروع إنجاز معجم عربي حاسوبي. ورغم طموح المشروع في جوانب عديدة نجد الأوراق البحثية المقدمة في المؤتمرين، اللذين جعلا لأوراق خبراء المعجم الحاسوبي، تهمل- أو تقصر- في معالجة الجوانب الدلالية للمعجم. وستعرض الورقة لذلك في موضعه.

(٢) يراجع تقرير المعجم العربي التفاعلي: http://www.almuajam.org/index2.htm

(3) **N. Calzolari :** Semantic web vision.

(٤) نهاد الموسى؛ العربية: نحو توصيف جديد في ضوء اللسانيات الحاسوبية. ص ٢٥٢

(٥) يرى حلام الجيلالي أن " قراءة أي مقدمة من مقدمات المعاجم المعاصرة، تجعلنا لا نقف على إشارة إلى المنهج المراد استثماره في تعريف المداخل [...] مما يجعل إعادة صياغة التعريف في المعاجم المعاصرة عملية نادرة أو غير معممة ": (تقنيات التعريف في المعاجم العربية المعاصرة) ص١١١. كما يرى محمد رشاد الحمزاوي أن من أهم أسباب التخبط في التطبيق المعجمي العربي " سذاجة المعجم في التهاون بقضايا شديدة التأثير، لاسيما قضية التعريف وتفاصيلها المتعددة ": (المعجم العربي: إشكالات ومقاربات)، ص ٢٨٦.

(٦) حسين نصار: المعجم العربي نشأته وتطوره، تمام حسان: اللغة العربية معناها ومبناها ص: ٣٢٠، محمد فتيح: في الفكر اللغوي ص ٢٧٠، محمد أحمد أبو الفرج: المعاجم العربية في ضوء علم اللغة الحديث، ص ١٧٨.

- غياب دراسة عربية- على حد علمنا- تتصدى لمعالجة المحتوى المعجمي الدلالي العربي تحليلا وتمثيلا، في صورة تعريفات معجمية.
- افتقار المجتمع الحاسوبي اللساني إلى مصادر معجمية عربية[٧] تواكب الطفرة في مجال الويب الدلاليSemantic Web، وسبب ذلك طريقة معالجة المحتوى المعجمي الدلالي العربي تحليلا وتمثيلا.
- عدم جدوى توظيف منهجية مقترحة في بيئة لغوية مختلفة عن العربية، إذ إن التعريف يحوي جل

المعرفة المعجمية بوصفه تمثيلا لمحتوى الوحدة المعجمي: الصرفي والتركيبي والدلالي. لذا فإن عملية التأسيس النظري لتعريفات معجم، وإجراءات ضبط محتواها ولغتها لابد أن ينبع من تصور نظري خاص بالطبيعة الصرفية والدلالية والتركيبية للغة المعجم؛ فلكل لغة طبيعتها الصرفية الدلالية التي تفرض طبيعة المعالجة النظرية لتعريفات معجمها. وتفرض اللغة العربية، بطبيعتها الصرفية الاشتقاقية، تأسيسا نظريا لتعريفات معجمها يتلاءم وطبيعتها من جهة، ويكون قابلا للتطبيق بفاعلية على مستوى المعجمين التقليدي، والحاسوبي من جهة أخرى.

جدوى الدراسة:

❖ على المستوى النظري:

إن فحص المعجم العربي في ضوء التصورات النظرية الغربية كفيل بأن يقدم خدمات للجانبين المعجميين الغربي والعربي: فنظرية مثل نظرية بوسطيوفسكي Pustejovsky (وهي نظرية تدعي العالمية universality ؛ لأنها تتحدث عن وسائط Parameters عوض الحديث عن قيم معينة Attributes ؛ مما يجعلها- في تصور أصحابها- صالحة للتطبيق على أي لغة؛ لتوجهها إلى الجانب الذهني أكثر من توجهها إلى التحققات الواقعية لمعجم بعينه. ولكن عند درس المعجم العربي بها ودرسها بالمعجم العربي وجدت فجوات في كلا التصورين؛ مما جعل الدراسة تقترح البنية التعريفية الصرفية[٨].

❖ على مستوى معالجة اللغات الطبيعية:

يساعد تعميق المحتوى المعجمي الدلالي وضبط لغته، في إنجاز كثير من التطبيقات اللسانية الحاسوبية، مثل فك اللبس المعجمي الدلالي[٩] Disambiguation ، والتوليد Generation، والتحليل Analysis.

أهداف الدراسة:

١. اقتراح منهجية لمعالجة المحتوى المعجمي الدلالي العربي تفي بمتطلبات المعالجة الآلية للعربية[١٠]، وتشمل:
   - جوانب التحليل، لاستخلاص المحتوى المعجمي الدلالي وتعميقه.
   - جوانب التمثيل، لصياغة معالم لغة للتعريف المعجمي، تكون منضبطة وقابلة للتعامل الحاسوبي.
٢. اقتراح تطبيق لساني حاسوبي يفيد من نتائج المنهجية المقترحة في تطوير المعجمية الحاسوبية (نظام محرر التعريفات العربية Arabic Lexical Definition Editor System).

المنهج وطبيعة المقاربة:

تتبنى الدراسة، في معالجتها المحتوى المعجمي الدلالي العربي، المنهج التوليدي Generative، حسب بوسطيوفسكي Pustejovsky. وينبغي هنا أن تقدم الدراسة احترازين:

- أنها أفادت من مناهج معجمية دلالية سابقة أخرى: في بعض فروضها ومقولاتها النظرية، وفي توظيف بعض إجراءاتها. ومن هذه المنهجيات المقاربة العلاقية Relational Approach، وإجراءات التحليل التكويني Componential Analysis، ووظفت كل ذا في إطار تصور خاص للمعجم العربي ومنهجية تحليله وتمثيله. كما أن الدراسة أفادت من المنجز المعجمي العربي (من خلال تحليل نصوص مدونة تعاريف المعجم الوسيط، والمعجم العربي الأساسي)، والمنجز العربي النقدي والتنظيري في مجالات المعجم والاشتقاق والدلالة.
- أنَّ تبني المنهج التوليدي في المقاربة لا يعني التطبيق الكامل لأفكاره وإجراءاته، بل يعني اقتناعا بفكرته الأساسية، وهي كيفية توصيف الوحدات المعجمية توصيفا يسمح برصد آليات تآلف الكلمات فيما بينها على المستوى الدلالي؛ للسيطرة على المعاني الإبداعية الجديدة (غير المنصوص عليها في المعاجم). فالاهتمام يكون بأنماط الدلالة الممكنة وآليات وقوعها أكثر من الاهتمام بالمعاني الكائنة والاعتناء بسردها في المعجم. غير أن الدراسة قد أضافت بنية أخرى على تصور بوسطيوفسكي لأبنية التمثيل المعجمي الدلالي وهي البنية الاشتقاقية، كما أعادت النظر في بنية التوارث لتقصيرها عن تمثيل البنية المعجمية الدلالية العربية.

---

(٧) أشار تقرير: "نملار" NEMLAR في دراسته المسحية للمصادر المعجمية العربية الحديثة، في إطار معالجة اللغات الطبيعية، إلى وجود نقص حاد في المصادر المعجمية العربية المعمقة دلاليا التي يمكن استثمارها في تطبيقات معالجة اللغة الطبيعية **N. L P.**

(٨) ظهرت الجوانب الصرفية في النظريات المعجمية الدلالية الغربية محكومة بطبيعة لغاتها وإشكالاتها، فقد خلا تصور بوسطيوفسكي من الجانب الصرفي تقريبا، كما أن معالجة ميلتشوك للجانب الصرفي في نموذجه محدود بإشكالات اللغة الفرنسية (اللواصق والتصريف).

(٩) تقصد الدراسة بمصطلح "المعجمي الدلالي" المحتوى الصرفيَّ الاشتقاقي الدلالي، والمحتوى التركيبي الدلالي، إضافة إلى المعنى.

(١٠) نصّت نتائج إحدى دراسات علاقة المحتوى المعجمي بمعالجة اللغات الطبيعية: "ينبغي أن تتضمن معالجة المستوى الدلالي في صناعة معجم محوسب على:إنشاء روابط دلالية جلية، ضمان الاتساق في محتوى المداخل، في لغة نصوص التعريف، في تقسيم المعاني":

Nancy Ide & Jean Véronis : Machine readable dictionaries: What have we learned, where do we go ?

- طبيعة المقاربة اللسانية: هي مقاربة قائمة على المدونة النصية Corpus based ، تستخلص معارفها اللغوية بتحليلها.
- أما عن طبيعة المقاربة اللسانية الحاسوبية: فالمجتمع اللساني الحاسوبي يعرف مقاربتين: إحصائية Statistical Based، وقواعدية Rule Based أو معرفية Knowledge Based . والدراسة تنتمي إلى النوع الثاني، إذ يذهب أصحاب هذا التوجه إلى أنه يمكن صياغة المعارف اللغوية أو تمثيلها في نماذج أو قوالب Modules منتظمة للحاسوب، بطريقة تحاكي الطريقة التي يتصرف بها الإنسان لغويا فهما وتوليدا. وهذه المقاربة تعتمد- أساسا- على بناء مصادر لغوية، معجمية أو نحوية، تحوي توصيفا مفصلا وجليا للغة.

أهم معالم التصور المقترح:

يفيد التصورُ المقترحُ من الأعمال المعجمية المنجزَة لتحقيق غايات البحث من تعميق التحليل للبنية المعجمية الدلالية وتدقيق لغة التمثيل، بما يجعله صالحا للتناول الحاسوبي. ويمكن إيجاز أهم معالم المنهجية المقترحة فيما يلي:

- يتوافق التصور المنهجي المقترح مع الفكر البرمجي والتطويري السائد أخيرا: فكر البرمجة الكائنية التوجه Object Oriented Programming، في قيامها على فكرة الفئة Class والكيان Object ، فالأول يحمل كل سمات النمط، والثاني يُشْتَق منه لكونه يعد واحدا من عدد لانهائي من تحققاته.

وقد ظهر هذا الجانب في الدراسة في افتراض نمطين للوحدات المعجمية الاشتقاقية: الأصول والفروع أو المشتقات المطردة، ومن ثم لتعريفاتهما، بافتراض نمطين من التعريفات : تعريفات أصلية للأصول وأخرى فرعية لمشتقاتها تُشتق منها.

- حاولت الدراسة الالتزام بمجموعة من الأفكار الحاسوبية التي تُكَوِّن الذهنية الحاسوبية أو الفكر الحاسوبي، مثل القول بالذرية في تحليل الوحدات المعجمية (السمات الدلالية)، وفي تمثيلها (الوحدة التعريفية Defimes)؛ لاتساق ذا مع أفكار التركيب والتوليد والقابلية للتحليل، والصياغة البنيوية، والتقليص في المعلومات بعدم الاضطرار إلى تكرارها، والقابلية لإعادة الاستخدام Reusability .

إجراءات التحليل ومدونته وأدواته:

اتخذت الدراسة مدونة لتطبيق مقولات التحليل عليها.

- **بنية المدونة:**

اعتمدت الدراسة مدونتين للتحليل:

- مدونة لغوية نصية Textual Corpus.
- ومدونة نصوص تعريفات معجمية، من المعجم الوسيط والمعجم العربي الأساسي.

وشملت مدونة النصوص حوالي ٨٠٠٠٠ ألف كلمة ضمّت نصوصا عربية فصيحة حديثة ومعاصرة، توزعت على مجالات مختلفة كالأدب والدين والسياسة والطب والرياضة. وقامت الدراسة بمعالجة مدونة فرعية Sub corpus من المدونة النصية العامة، هي الكلمات التي يبدأ جذرها بحرف الجيم، باعتبارها نموذجا للمدونة النصية الكبرى، كما عالجت مدونة من نصوص التعريفات المعجمية العربية.

التحليل وأدواته:

اتخذت الدراسة أربعة مستويات للتحليل، تتآزر لتحليل المحتوى المعجمي الدلالي لوحدات المعجم، وهذه المستويات هي:

- المستوى الصرفي(الاشتقاقي): وشمل تحليل العلاقات الاشتقاقية الدلالية، واقتراح اتخاذ أصل الاشتقاق كلمة متحققة لا مجرد جذر، وتبع ذلك اقتراح لإعادة تنظيم البنية الصغرى Microstructure للمعجم العربي.
- والمستوى التركيبي: ويشمل الوسم التركيبي الذي يحدد قسم الكلام الذي تنتمي إليه الوحدة المعجمية، وتوضيح طبيعة البنية المعجمية للحدث Event، وتوضيح القيود الانتقائية، وطبيعة المكملات التي تحتاج الوحدة المعجمية المعرّفة إليها.
- والمستوى الدلالي: ويشمل توضيح الحقل الدلالي للوحدة المعجمية، والعلاقات الدلالية التي تشترك فيها، واستخلاص السمات الدلالية.
- والمستوى المعجمي: ويشمل تحديد أنماط الوحدات المعجمية بوصفها معرَّفات.

وقد استخدمت الدراسات أدوات حاسوبية للتحليل هي المحلل الصرفي Morphological Analyzer والمفهرس الآلي Concordancer، وبرنامج لاستخلاص التصاحبات اللفظية N-gram إضافة إلى التحليل اليدوي.

جوانب التمثيل (لغة التعريف):

تتكون لغة التعريف - قياسا على اللغة الطبيعية – من معجم لوحدات التعريف Defimes ، ونحو لجملة التعريف Definition Syntax ، ومجموعة من أبنية التعريف وتقنياته.

فوضعت الدراسة معالم **معجم لوحدات التعريف** Defimes ، يمكن أن يستخدم في تعريف مفردات المعجم.

ونحو التعريف Definition Syntax يعني اقتراح أنماط تركيبية لجملة التعريف بحيث تحقق غايتين، الأولى: تحقيق الاتساق في التحرير، والثانية هي أن يكون نص التعريف قابلا للتحليل الجملي الآلي Parsing.

وأبنية التعريف Definition Structure هي مكونات التعريف الفرعية، وتتوازى مع مستويات التحليل. وخرجت الدراسة بتصور لبنية التعريف اعتَبَرَتْ فيه صياغةَ البنية المعجمية الدلاليةLexical Semantic Structure موضوعا مركزيا في التمثيل، فتتبعت عناصر صياغتها على المستويات المختلفة للتحليل والتمثيل.

أولا: أهم مقومات تصور نظري لتحليل المحتوى المعجمي وتمثيله:

ينبغي- قبل تناول معالم تصور الدراسة للمنهجية المقترحة - البدءُ بتوضيح أهم مقومات أي تصور نظري يهدف إلى معالجة المحتوى المعجمي الدلالي تحليلا وتمثيلا، وتحديد شروط وجوده، وأهم متطلبات بناء منهجيته. وتقترح الدراسة شروطا ينبغي أن تتوفر في أي تصور نظري يحاول أن يقارب المحتوى المعجمي الدلالي بغرض تمثيله في صورة تعريفات معجمية. فينبغي أن تتوافر فيه المقومات الآتية:

- تصور شامل لمنظومة اللغة، ومنظوماتها المكونة، وموقعية المعجم بينها.
- تصور واضح لبنية معجم اللغة: كياناتٍ وعلاقات، وآليات؛ لكيفية صياغة هذه البنية وتحقيقها في المتن المعجمي.
- تصور لطبيعة المحتوى المعجمي الدلالي (الصرفي والدلالي والتركيبي) والعلاقة بين مكونات هذا المحتوى، وكيفية تعاونها في تمثيل المحتوى• وتصور لطبيعة المعنى.
- تصور لمنهجية التحليل: منطلقات نظرية وإجراءات تطبيقية، ووحدات التحليل لكل من مستويات اللغة.
- تصنيف لأنماط المعرفات، باعتبارها أولى خطوات تنظيم مادة المعجم، وتصنيف كياناته وأنماطه المَعرِفية.
- تحديد لمتطلبات كل صنف من أنماط المعرفات، وطبيعة كل نمط، وسماته الشكلية والمضمونية.
- إجراءات لتحليل المحتوى: سواء أكانت إجراءات يدوية أو إجراءات حاسوبية، تكون متسقة مع المنطلقات النظرية من جهة، وفعالة منتجة من جهة أخرى.
- استراتيجيات للتمثيل: فالمعلومات المعجمية الدلالية ليست صنفا واحدا بل أصناف متنوعة متباينة، كما أن الوحدات المعجمية نفسها أنماط متباينة، لكل نمط منها خصائصه.
- معجم للتمثيل: يشتمل على الوحدات التي يمكن تمثيل الوحدة المعجمية بها.
- قوالب أبنية تركيبية تمثيلية (مستوعبة، ممثلة، نسقية).
- رؤية واضحة لحدود التصور وأهدافه، وموقفه من التطبيقات الحاسوبية.
- آليات تكون الكلمة/ الوحدة المعجمية: الاشتقاق (وأثره في بناء المعجم والبنية الصغرى)، النحت، الاقتراض، الاصطلاح.
- الاطرادات الصرفية والدلالية، وكيفية توظيفها في تقليص حجم المعجم Dictionary.
- تصور لوحدة التحليل في كل مستوى.
- تصور لوحدة التمثيل.
- تصور لبنية التعريف ومكوناته الجزئية.

أولا: المنطلقات النظرية: الصرفية والتركيبية والدلالية:

تعددت التصورات النظرية التي قاربت الدلالة المعجمية بغرض تحليلها وتمثيلها، وهي في تعددها لم تبدأ واحدة منها من فراغ، بل أفاد كل تصور لاحق من التصورات السابقة. وقد أفاد التصور النظري المقترح من تصورات نظرية متعددة؛ بغرض توظيفها في إطار التصور، مع المحافظة على الاتساق الداخلي للتصور. فالتصور متأثر بالتوجه التفكيكي بدءا من كاتز وفودر في تصورهما عن المدخل المعجمي(١١)، وأفكار تفتيت المعنى لدى بولينجر(١٢)، وأفكار تحليل المكونات لدى يوجين نايدا(١٣).

١. المنطلقات النظرية:

مفهوم الدراسة للتعريف: هو توصيف المحتوى المعجمي الدلالي للوحدة المعجمية في ذاتها، وفي علاقاتها داخل بنية المعجم.

البنية المعجمية الدلالية: هي إحدى غايات التصور النظري، ويمكن تعرفها في مرحلة التحليل عن طريق التوصل إلى الأنماط التي تحكم اللغة، مثل: أنماط الوحدات المعجمية، وأنماط الصيغ الصرفية. كما يمكن صياغتها على مستوى التمثيل عن طريق مجموعة من التقنيات، هي:

- **تنميط المُعَرَّفات:** أول عمل ينبغي أن ينجز في سبيل تنميط التعريفات، والصياغة البنيوية للتعريف.
- **فصل المشترك اللفظي** Homonyms بعد تمييزه عن متعدد المعنىPolysemy.
- **تصنيف الوحدات المعجمية** في حقول دلالية.
- **تجلية العلاقات الدلالية:** الأفقية (الترادف والتضاد)، والرأسية أو الهرمية (الاشتمال، والتضمن).

(١١) Kats & Fodor, (1963), the structure of semantic theory .P: 178.

(١٢) Dwight Bolinger, The Atomization of Meaning. P:22.
http://www.jstor.org/pss/411524

(١٣) Nida, Eugene, A Componential Analysis of Meaning. An introduction to semantic structures p: 13.

٢. في طبيعة المحتوى المعجمي الدلالي:

❖ **الصرف:** تعد الصيغة الصرفية البنية التحتية للمحتوى الدلالي، كما أن للمقولة الصرفية دورا كبيرا في تحديد المعنى. وفي مستوى الصرف اعتمدت الدراسة الأفكار التالية:

- **الوحدات المعجمية** نوعان: اشتقاقية لها عمق صرفي دلالي، وأخرى غير اشتقاقية لا تحوي دلالة صرفية.
- **الوحدات المعجمية الاشتقاقية** نوعان: أصول وهي الوحدات المعجمية الأبسط والتي يشتق منها وحدات معجمية أخرى، وفروع وهي الوحدات المعجمية التي تشتق من الأصل.
- إعادة النظر في **بنية المداخل المعجمية العربية**، وبالتالي طريقة تنظيم بنية المعجم الكبرى.
- **الجذر** فرضية تصنيفية لا تفسيرية.
- **الصيغة الصرفية** (للوحدات الاشتقاقية) هي الوحدة الدلالية الصرفية الصغرى.
- **المشتقات الاسمية** صنفان: مشتقات مطردة، ومشتقات متحولة معجميا.
- **المقولة الصرفية** هي الوظيفة الصرفية الدلالية (مثل اسم الفاعل، اسم المفعول... الخ).

❖ **الدلالة:**

هي مركز المعالجة المعجمية الدلالية. وقد أفادت الدراسة من مقولات الحقول الدلالية، وتحليل المكونات، وتصورات التوجه العلاقي، في بيان البنية المعجمية الدلالية عن طريق العلاقات الدلالية.

❖ **التركيب:**

يرتبط التركيب بالدلالة ارتباطا شديدا؛ نظرا لتوقف بعض الدلالات على المحتوى التركيبي، والعكس صحيح: فالدلالة تتحكم في النمط التركيبي الممكن للوحدة المعجمية. واتخذت الدراسة في هذا الجانب توجها سياقيا أفقيا Syntagmatic، فحاولت تنميط القيود الانتقائية، واستقصت أنماط بنية الحدث للفعل.

ثانيا: التحليل : مدونته، وإجراءاته، ومخرجاته.

مدونة الدراسة:

المدونة اللغوية Corpus[١٤] مجموعة نصوص متجانسة، تُجمع على أسس معينة يحكمها غاية المدونة نفسها واستعمالها، وتطبيقاتها المحتملة. تنظم في شكل معين (شكل حاسوبي في الغالب)، تكون مادة للتحليل اللغوي (أو الثقافي والحضاري والنفسي)، واستخلاص الظواهر حاسوبيا.

والمدونة النصية لا تقدم معلومات مباشرة عن اللغة، ولكن التحليل (اليدوي أو الحاسوبي) هو الذي يستخلص هذه المعلومات؛ لذا يستعين العاملون في مجال المدونات الحاسوبية بأدوات حاسوبية تقوم بإجراءات التحليل: من تصنيف وإحصاء وتصنيف للمادة. والمدونات المحوسبة من حيث المعالجة نوعان: مدونة ذات تعليقات (Annotated)، وأخرى خام Raw.

ولقد غدت المدونات اللغوية المحوسبة، بعد الثورة اللسانية الحاسوبية، إحدى العلامات المميزة للسانيات عموما، واللسانيات الحاسوبية على وجه الخصوص؛ حتى ظهر مجال لساني جديد أُطلِقَ عليه لسانيات المدونة Corpus Linguistics، " لقد حَوَّل الحاسوب المدونةَ النصيةَ (التقليدية) من مجرد الاهتمام بموضوع بحثي تخصصي واحد إلى كونها أصبحت مفتوحة على كل الموضوعات إلى أن أحدثت المدونات المحوسبة ثورة في دراسة اللغة "[١٥].

والمدونة في أحد تصوراتها هي اللغة من خلال نماذج ممثلة لها، لذا يحرص المتخصصون على اختيار نصوصها، وتصميم قاعدة بياناتها. فالمدونة النصية يمكن دراسة مستويات اللغة من خلالها، كما يمكن دراسة ظاهرة أو موضوع من خلال مادتها. وقد ارتبطت المدونات النصية المحوسبة بصناعة المعاجم؛ فلم يعد من المستساغ (في الغرب) الحديث عن معجم دون البدء بالحديث عن مدونته النصية.

تصميم المدونة النصية:

تنوعت مادة المدونة من حيث المجال: (أدب، دين، سياسة واقتصاد، إنسانيات، فنون، رياضة، علوم وتقنية، جغرافيا وفلك)، ومن حيث الجنس الأدبي للمادة الأدبية: (رواية، قصة، مقالة)، ومن حيث تخصص الكُتَّاب (رجال، إناث)، ومن حيث التوزيع الجغرافي (مصر، الجزائر، المغرب، سوريا)، ومن حيث طبيعة الوسيط الإعلامي: (كتب، صحف، مجلات، مواقع إلكترونية)، ومن حيث طبيعة النص (عربي، مترجم إلى العربية).

وصف المدونة النصية:

ليس من غاية الدراسة إنشاء مدونة؛ فهذا موضوع يخرج عن حدود موضوعها وأهدافها، غير أن اعتمادها على المدونة أمر ضروري لازم على المستوى المنهجي، فكان من الممكن- والأمر كذلك- أن تعتمد الدراسة إحدى المدونات العربية المنجزة بالفعل، وتُجري عليها تحليلاتها. غير أن الواقع العملي يتناقض مع ذلك: فجل من يمتلكون مدونات لا يتيحونها للاستعمال العام، وما يتاح منها يتاح بمقابل مالي باهظ، مع عدم قيامه على أسس تتوافق مع منهجية البحث.

---

(١٤) تعددت المقابلات العربية لهذا المصطلح الإنجليزي ومنها المكنز، والذخيرة اللغوية، والمدونة الذي اعتمدته الدراسة، وقد اعتمدته لكونه المقابل العربي الأكثر شيوعا واستعمالا في أوساط اللسانيين المختصين، وكأنه نوع من الإجماع: (فقد عقد أكثر من مؤتمر في تونس والمغرب تحت هذا العنوان، ونوقشت دراسات تحمل مصطلح المدونة عنوانا، في مصر والمغرب العربي أيضا). إضافة إلى أن المصطلحات الأخرى المستخدمة لها دلالات أخرى في حقل اللسانيات، فمثلا مصطلح مكنز قد يقابل (Corpus) وقد يقابل (Thesaurus).

(١٥) Susan Hunston: Corpora in Applied Linguistics p:1

على أن قيمة أي مدونة تعود، بعد دقة تصميمها وعمق تأسيسها اللساني، إلى التحليلات المجراة عليها: من وسم تركيبي Part of speech tagging ووسم دلالي Sense tagging ، وتعليق وإضافات Annotation صرفية دلالية وتركيبية.
والدراسة الحالية لا تحتاج لمدونة محشاة، بل تحتاج إلى مدونة غير معالجة Raw Corpus ، فأنشئت مدونة نصية (مبسطة) تؤدي المهام التي تفرضها إجراءات الدراسة في التحليل والتمثيل.
ويمكن توصيف المدونة بأنها مدونة للغة العربية الفصيحة في العصر الحديث، من خلال مجموعة من الأعمال الأدبية والدراسات العلمية والإنسانية. حاولت الدراسة أن تكون ممثلة للغة العربية في الاستعمال الطبيعي العام؛ فقد تنوعت بين الكتب والمقالات. بين التأليف والترجمة.

	كلمات	مداخل
حجم المدونة	٨٠٦٣٣	٤١٤٥١
عدد المجالات	٨	
التغطية الزمنية	من ١٩٥٠ وما بعدها	

ويمكن أن نقدم مخططا عاما لمدونة الدراسة في الجدول التالي:

المجال	المجال الفرعي	النص	صاحبه
الأدب	رواية وقصة قصيرة	أصداء السيرة الذاتية	نجيب محفوظ (مصر)
		جلنار	زكرياء أبو ماريا (المغرب)
		العمامة والقبعة	صنع الله إبراهيم (مصر)
		ذاكرة الجسد	أحلام مستغانمي (الجزائر)
	دراسات أدبية	في الشعر الجاهلي	طه حسين (مصر)
		فصول في علم الاقتصاد الأدبي	حنا عبود
		تمظهرات التشكّل السير ذاتي قراءة في تجربة محمد القيسي السير ذاتيّة	محمد الباردي
الدين		جَدِّدْ حياتَك	محمد الغزالي
		اَلتَّصْوير اَلْفَنِّيّ فِي اَلْقُرْآن	سيد قطب
		مَعَانِي اَلْقُرْآن بَيْن اَلرِّوَايَة وَالدِّرَايَة	أَحْمَد حَسَن اَلْبَاقُورِيّ
اقتصاد وسياسة		مقالات	
		تاريخ الجزائر المعاصر دراسة (الجزء الأول)	محمد العربي الزبيري
طب وصيدلة		مقالات	
رياضة		مقالات	
فنون	سينما	مقالات	
	مسرح	مقالات	
	موسيقى	مقالات	
علوم إنسانية	انثروبولوجي (إناسة)	مـدخـل إلـى علم الإنسـان (الأنثروبولوجيا)	عيسى الشمّاس
	علم نفس	فـن الإصـغاء	إريك فروم ترجمة: محمود منقذ الهاشمي
علوم وتقنية		مقالات	
علوم طبيعية	جغرافيا	مقالات	
	فلك	مقالات	

مدونات التعريفات المعجمية:

حاولت الدراسة الإفادة من المنجز العربي في مجال التعريفات المعجمية، فاتخذت مدونة أخرى هي نصوص تعريفات المعجم الوسيط، وباب الجيم في كل من المعجم الوسيط والمعجم العربي الأساسي[١٦].

معالجة المدونة النصية:

(١٦) كان التعامل الإحصائي مع متن الوسيط كاملا بغرض استكشاف لغة الوسيط في التعريف. في حين كان التركيز على باب الجيم في المعجمين (الوسيط والأساسي) باعتبارهما عينة الدراسة التي عليها التركيز في اختيار أمثلة للتطبيق عليها في مرحلة التمثيل.

المدونةُ – أي مدونة - قبلَ المعالجة إنْ هي إلا مجموعة من النصوص المختارة، والمعالجة هي التي تجعل منها مدونة لغوية صالحة للإفادة منها في حقول اللسانيات.
ومعالجة المدونة تشمل جانبين: حاسوبي ولغوي، أما الحاسوبي فهو تهيئة البرامج التي تحوى متن المدونة تخزينا وتعاملا معها. وأما الجانب اللغوي فيشمل التحليلات على المستويات الصرفية والتركيبية والدلالية والاستعمالية. ويمكن إيجاز الإجراءات التي تعاملت الدراسةُ بها مع مدونتها النصية فيما يلي:

- تحليلات **N-gram** وهي تحليلات آلية لتحديد درجة تكرار الكلمةUni gram، أو العِبارة Tri gram & bi gram ، أو التصاحبات اللفظية Collocations في المدونة، مع عرض النتائج مرتبة حسب عدد مرات التكرار، وذكر النسبة إلى عدد كلمات المدونة.
- استخلاص مدونة فرعية sub corpus (الوحدات المعجمية التي تنتمي إلى الباب المعتمد عينة في الدراسة وهو باب الجيم).
- التحليل الصرفي Morphological Analysis (وناتج هذه المرحلة وحدات معجمية مصنفة حسب المقولة الصرفية بالإضافة إلى سياقاتها) أو Morphological clustering .
- التحليل الدلالي.
- التحليل التركيبي.
- التحليل المعجمي.
- تحديد رؤوس المداخل Lemmatization .
- الربط بين مدونة تعريفات الوسيط والأساسي ومداخل المدونة النصية الفرعية.

Matching between lexical entries in dictionaries and sub corpus

معالجة مدونة التعريفات:

- التحليل الإحصائي للغة التعريف (كلمات وتصاحبات لفظية).
- التحليل الصرفي والدلالي.
- الربط بالمدونة النصية.
- الاستعانة بها في الوسم الدلالي والتركيبي في المدونة النصية.
- التوثيق.

منهجية الدراسة

طبيعة المعالجة:

- تنتمي الدراسة- منهجيا- إلى المقاربة القواعدية Rules Approach ، وهي مقاربة معرفية الأساس Knowledge based ، استعانت ببعض جوانب الإحصاء الحاسوبي بوصفها إجراءات تخدم التحليل(١٧).
- كما يمكن تصنيف الدراسة باعتبارها دراسة معتمدة على المدونة النصية(١٨) Corpus-based approach.
- أما في التحليل الدلالي فقد أفادت الدراسة من التحليل السياقي الأفقيSyntagmatic Analysis ويعتمد على التحليل السياقي للتوصل إلى المحتوى الدلالي، والتحليل الرأسي الاستبدالي Paradigmatic Analysis في تحليل العلاقات الدلالية والشبكة الصيغية.
- وتنتمي الدراسة إلى الدلالة المعجمية الحاسوبية، حيث تقترب في جانب التحليل من علم الدلالة المعجمية، كما تنتمي في جانب التمثيل إلى المعجمية الحاسوبية.

مستويات التحليل:

قدمت الدراسة تحليلا في أربعة مستويات هي:

- المستوى الصرفي. - المستوى التركيبي.
- المستوى الدلالي. - المستوى المعجمي.

الإجراءات:

إجراءات التحليل هي المعالجة اليدوية أو الحاسوبية للمدونة استكشافا للظواهر، وتحليلا لها.
وقد كان للدراسة توجهان في التحليل: الأول تحليل يشمل المدونة كلها، ويكون تحليلا حاسوبيا غايته مرحلة التحليل، والثاني تحليل يقتصر على باب الجيم (المدونة الفرعية لكلمات الجيم)؛ لأغراض التحليل المعمق بغرض التمثيل ، ويكون تحليلا يدويا في الغالب.

إجراءات التحليل الدلالي:

- **الوسم الدلالي** Sense Tagging: ويعني تحديد قسم الكلام التي تنتمي إليه الوحدة المعجمية.

---

(١٧) توجد مقاربتان في حقل اللسانيات الحاسوبية: مقاربة قواعدية تهدف إلى وضع قواعد لسانية للحاسوب، ومقاربة إحصائية تترك للحاسوب استكشاف الاطرادات الموجودة في المصدر اللغوي أو المدونة المحوسبة، وصنع قواعده.

(١٨) تقسم المقاربات اللسانية الحاسوبية من هذه الناحية مقاربتين: الأولى معتمدة على المدونة (Corpus-based approach)، والثانية مقاربة لسانية تعتمد المنجز اللساني بالأساس في استخلاص المعارف اللغوية.

- **التحشية الدلالية** Sense Annotation : أي تسجيل المعاني المستنتجة من خلال السياق، وتدوين الحواشي الدلالية مثل العلاقات الدلالية، والحقول الدلالية، والسمات الدلالية.
- **الترابطات الدلالية** Sense Clustering ويقصد به تجميع المعاني المترابطة للوحدة المعجمية.

إجراءات التحليل الصرفي:

- **التحليل إلى جذوع** Stemming : وهو إجراء غايته التوصل إلى الجذوع Stems أو الوحدات المعجمية الأبسط في المدونة مفهرسة، بحيث تصبح المخرجات قائمة من المداخل.
- **تحليل صيغ الوحدات المعجمية الاشتقاقية** Patterns Analysis : ويعني محاولة الوسم الصرفي الدلالي للصيغ الصرفية Morphological Patterns ، وتحديد الدلالة الصرفية لها، وتحديد المقولة الصرفية للمشتقات.

إجراءات التحليل التركيبي:

- **الوسم المقولي التركيبي** POS tagging : وهو الوسم التركيبي للوحدات المعجمية من حيث مقولتها التركيبية، وانتماؤها إلى أحد أقسام الكلام (الفعل، المشتق، الاسم، الأداة).
- **تعيين فئات الأفعال** Verb classes.
- **تصنيف الفعل حسب نمط بنية الحدث** Event Structure.
- **الأنماط التركيبية للاسم المركب**.

إجراءات التحليل المعجمي:

- **تعيين رؤوس المداخل** Lemmatization: ويعني تجميع التصريفات المترابطة، وتحديد رأس لها head word يتخذ عنوانا لها في المعجم، أو في قاعدة البيانات.
- **فصل المشترك اللفظي** Homonymy distinctions.
- **التصنيف حسب المقولة المعجمية** (كلمة، تعبير اصطلاحي، فعل عباري ...).

الأدوات الحاسوبية لتحليل المدونة:

الأدوات الحاسوبية التي استخدمتها الدراسة في التحليل أدوات جاهزة هي:

المفهرس الآلي Concordancer [١٩]:

يُعَد المفهرس الآلي واحدا من الأدوات الأساسية الأكثر أهمية في التعامل مع المدونات الحاسوبية؛ إذ تعتمد عليه كثير من المشروعات المعجمية في تحليل مدونتها. فهو يقوم بتصنيف الكلمات، وفهرستها، وتحديد مرات التكرار، ونسبته مقارنا بعدد كلمات المدونة كلها، وتحديد التصاحبات اللفظية، وعرض المخرجات، أي أنه يقوم بتجهيز المدونة للمراحل التالية من التحليل. والمفهرس الذي اعتمدت عليه الدراسة هو: "Conc App".

خوارزم التصاحبات (N-Gram):

وهو عبارة عن خوارزم Algorithm لتحديد التصاحبات اللفظية Collocations جُعلت له واجهة للمستخدم User Interface[٢٠]، وهو يستكشف الكلمات التي تقع متصاحبة، ويرتبها تصاعديا أو تنازليا، مع ذكر نسبة ورود كل تكرار. وقد استخدم في تحديد التصاحبات في المدونة النصية، وفي مدونة التعريفات.

وإلى جانب اعتماد الدراسة في التحليل على أدوات حاسوبية، اعتمدت أيضا التحليل اليدوي في كثير من جوانب التحليل؛ وذلك إما لتحقيق قدر أكبر من الدقة، أو لعدم توفر الأداة الحاسوبية التي تصلح لذلك.

وفي الشكل التالي تستعرض الدراسة جوانب منهجية العمل على مستوى التحليل والتمثيل[٢١]:

الموضوع	التحليل	التمثيل
الغاية	التوصل إلى العنصر الذري في: الوحدات المعجمية. المعاني.	صياغة المعنى الذري (ناتج التحليل) في صورة تعريف معجمي.
المدخلات	مدونة نصية محوسبة. مدونة للتعريفات المعجمية العربية مصادر التوثيق	وحدات معجمية مصنفة. معانٍ.
المخرجات	وحدات معجمية مصنفة. معانٍ. سمات دلالية.	تصنيفات. أبنية تعريفية.

(١٩) رابط المفهرس المعتمد في التحليل: **http://www.edict.com.hk/PUB/concapp/**

واسم الشركة المنتجة له: Educational Services Consultants - والنسخة المعتمدة : Version = 4.00.000

(٢٠) الأداة من إعداد المبرمج المصري: حاتم مصطفى .وهي موجودة على الشبكة الدولية للمعلومات من خلال الرابط التالي:-**http://www.n-gram\CodeProjectN-gramandFastPatternExtractionAlgorithm**

(٢١) هذه الإجراءات لها جانبان، فالأول بكونه إجراءات الدراسة للتحليل، والثاني بكونه جزءا من التصور المقترح، وهو الجزء المتعلق بإجراءات التحليل.

الإجراءات	- التحليل اليدوي. - التحليل الحاسوبي. - التحليل السياقي. - تحليل التصاحبات.	تنميط المعرفات. تصميم بنية التعريف ومكوناتها. صياغة معجم مفردات التعريف. قواعد التعريفات.
المستويات	المستوى الصرفي. المستوى التركيبي. المستوى الدلالي. المستوى المعجمي. مدونة التعريفات المنجزة.	أنماط المعرفات. معجم مفردات التعريف. أنماط التعريفات بنية التعريف Definition Syntax البنية المعجمية الدلالية.
المحتوى الصرفي	مشتق / غير مشتق. الأصول والفروع. المقولة الصرفية. مشتق مطرد/ مشتق متحول معجميا	البنية التعريفية الصرفية: تعريف الأصول، والفروع. تعريف المشتق المطرد. تعريف المتحول معجميا. تمثيل محتوى الصيغة الصرفية.
المحتوى التركيبي	المقولة التركيبية. البنية المحمولية.	البنية التعريفية التركيبية: التصنيف.
المحتوى الدلالي	الحقل الدلالي. البنية العلاقية الدلالية. السمات الدلالية. نمط بنية الحدث.	البنية التعريفية الدلالية: صياغة الحقول الدلالية. التصنيف. السمات الدلالية. جملة التعريف.
المحتوى المعجمي	المقولة المعجمية. الوحدة المعجمية.	أنماط المعرفات.
المنطلقات النظرية	الجذر وإشكالية أصل الاشتقاق. المعنى الصرفي: طبيعته. الصيغة الصرفية وحدة دلالية.	طبيعة المحتوى المعجمي الدلالي. أبنية التعريف. معجم وحدات التعريف (الديفيمات). اشتقاق تعريفات المشتقات من تعريفات الأصول البنية المعجمية الدلالية.
	المقولة التركيبية. البنية الحملية. القيود الانتقائية.	أقسام الكلم. الأنماط التركيبية للفعل. أنماط القيود الانتقائية.
	السمات الدلالية. العلاقات الدلالية الأفقية، والرأسية. القيود الانتقائية.	صياغة قائمة السمات الدلالية ونحوها. صياغة العلاقات.
	الوحدة المعجمية العامة، والموسوعية والمصطلحية. الوحدة المعجمية البسيطة، ومتعددة الكلمات	بنية الوحدة العامة. بنية الوحدة الموسوعية. بنية الوحدة المصطلحية.
المشترك اللفظي	تحديد المعايير المميزة، والتي تفصله عن متعدد المعنى. الفصل العملي. تحديد قائمة المداخل النهائية للمعجم.	اتخاذ كل وحدة رأسا لمدخل معجمي.
متعدد المعنى	فصله عن المشترك اللفظي. استخلاص المعاني المستقلة. التعامل مع المعنى الواحد باعتباره وحدة المعجم الدلالية.	اتخاذ موقف من تنظيم المعاني لمتعدد الدلالة. تبين العلاقات الدلالية داخل المدخل الواحد - لكل معنى تعريف.
العلاقات الدلالية	نمطان من العلاقات الدلالية: رأسية أفقية	صياغتها في جملة التعريف، وفي البنية الدلالية للتعريف.
تنميط المعرفات	تحديد أنماط المعرفات.	تحديد متطلبات كل نمط.
البنية المعجمية	تحديد متطلباتها	صياغتها على المستويين اللغوي والحاسوبي

ثالثا : التمثيل : لغته، وأبنيته.

لغة التعريف: المقصود بلغة التعريف اللغة المستخدمة في تمثيل الوحدة المعجمية. واللغة هنايقصد بها الوحدات التعريفية والعبارات والجمل وقواعد التركيب.

- **معجم وحدات التعريف:** ويعني قائمة الكلمات المستخدمة في تحرير التعريف، وهي مختارة على أسس من نتائج التحليل الصرفي والتركيبي والدلالي والمعجمي.
- **نحو التعريف:** هو قواعد تآلف الوحدات التعريفية في جملة التعريف، بحيث يمكن تحليل جملة التعريف (آليا) إلى مكونات يمكن الإفادة منها دلاليا في التعرف البنية الدلالية.
- **محددات التعريف:** هي كلمات أو عبارات أو تصنيفات تأتي مع تعريف وحدة معجمية؛ كي تميزه عن تعريف آخر لوحدة معجمية ذات دلالة متقاربة أو متداخلة معها.
- **بنية التعريف:** هي بنية نموذجية للتعريف ، تعد بمثابة قالب يقوم المحررون بالبناء عليه ، واستكمال محتواه لكل وحدة معجمية.

الوحدة التعريفية (الديفيم DEFIME)

المفهوم، الفروض، الطبيعة، الأصول، الإجراءات، الكفاية، قواعد التآلف:

في الوقت الذي طغى فيه التفكيك على مقاربة الفعل والمشتقات ومحاولة تمثيله دلاليا، أو دلاليا تركيبيا، لم توجد في المقابل جهودٌ تحاول تمثيل هذه المقاربات أو نتائجها على المستوى المعجمي الخالص في صورة تعريفات معجمية، تفيد من إجراءات التفكيك والتذرية ونتائجها.

١. **المفهوم**[٢٢]: هو الوحدة الصغرى التي تمثل الذرة البنائية التي يتكون منها التعريف، وهي وحدة تنتمي إلى لغة التعريف (التي هي لغة شارحة، تصف اللغة meta language) يمكن بناء التعريف بها، ويمكن تفكيكه إليها).

فالديفيم: نظرية في لغة التعريف، غير أن هذه اللغة- بوصفها تمثيلا للمعنى أو المحتوى- تقتضي الدراسة المفصلة للمحتوى؛ للتوصل إلى أولياته التي تتطلب اصطناع أوليات (ديفيمات) تعريفية توازيها. وتقوم الفكرة على افتراض مفاده: إن المعنى بنية مركبة أو معقدة ينبغي تحليلها أو تفكيكها لإعادة توصيفها وتمثيلها ديفيميا، إذ "هناك تصور دقيق مؤداه أن كل لفظة معجمية عبارة عن حزمة من الملامح [٢٣].

٢. **الفروض والمنطلقات** (التي تقوم عليها فكرة الديفيمات):

يمكن افتراض مجموعة منتهية من الوحدات الدلالية يبنى منها التعريف المعجمي.

- تقوم الفكرة على إمكان تذرية لغة التعريف في مقابل" تذرية المعنى"[٢٤]. ويمكن صياغتها كما يلي:
  - المعنى مركب من عدة أوليات.
  - يمكن اصطناع لغة واصفة Meta language للتعريف المعجمي، تعبر- بالتوازي- عن أوليات تمثل تلك الأوليات.
  - ينتج تمثيل المعنى في صورة تعريف من تراكب هذه الأوليات.
  - لتراكب هذه الأوليات قواعد rules of syntax تحكم تآلفها في بنية التعريف.

٣. **الطبيعة: خصائص:**

- تنتمي الوحدة التعريفية إلى اللغة الواصفة Meta Language المعجمية ولكنها وثيقة الصلة بتحليل اللغة / المعجم ؛ استخلاصا وتمثيلا.
- وحدة بناء التعريف المعجمي ليست دائما كلمة بل قد تكون عبارة أو تركيبا أو جملة.
- للوحدة التعريفية علاقة وثيقة ببنية المعجم التصورية Conceptual Structure [٢٥]؛ إذ هي تمثلها باحتوائها الوحدات المعجمية التي تكاد تمثل نموذجا لمعجم مصغر للغة.
- لها بنية (شكل) ودلالة (محتوى).
- يمكن التعامل معها تحليلا وتركيبا.

٤. **الأصول (الأصول الإبيستيمولوجية):**

ظهر مصطلح الفونيم Phoneme في أدبيات الدرس الصوتي، (علم وظائف الأصوات Phonology)[٢٦] وكان لظهور هذا المفهوم أثر في إحداث ثورة في مجالات البحث اللغوي كلها، باعتبار الفونيم وحدةً للتحليل يمكنها أن تقدم صياغة بنيوية أكثر كفاءة للبنية الصوتية للغة في علاقتها بالجانب الدلالي.

---

(٢٢) التعريف: تتميم الوصف الدلالي والتركيبي والصرفي للكلمة، أو التعبير المعرف، لمتطلبات المستخدم البشري ومتطلبات برامج الحاسوب.

(٢٣) جوديث جرين: علم اللغة النفسي، تشومسكي وعلم النفس. ترجمة وتعليق مصطفى التوني ص٨٤ .

(٢٤) وردت فكرة تذرية المعنى في ورقة بولينجر Dwight Bolinger: The Atomization of Meaning

(٢٥) مُقترِح المفهوم هو جاكندوف (Ray Jackendoff) ويذهب إلى أن البنية اللغوية جزء من بنية كبرى هي البنية المفهومية التي تجتمع فيها البنى الإدراكية للعالم.

ثم ظهر في مستويات التحليل اللغوي مصطلحات تعبر عن مثل هذا المفهوم فيما يخص طبيعة مادتها. فظهر في مجال علم الصرف مفهوم الوحدة الصرفية (المورفيم morpheme). وتتابع ظهور المفاهيم والمصطلحات المماثلة في بقية فروع التحليل اللغوي.

ويمكن **تحديد المصطلحات في مستويات اللغة** من خلال الجدول التالي:

المصطلح	المفهـــوم	المجال
Phoneme	الوحدة الصوتية الدنيا التي يؤثر تغيرها في تغير معنى الكلمة التي تدخل في تكوينها	علم وظائف الأصوات Phonology
Morpheme	الوحدة الصرفية الدنيا التي تحمل معنى، ويمكنها الدخول في تكوين كلمة أو الدخول على كلمة	علم الصرف Morphology
Lexeme	الوحدة المعجمية الدنيا	المعجمية Lexicology
sememe	الوحدة الدلالية الدنيا	علم الدلالة Semantics
synteme	الوحدة التركيبية الدنيا	التركيب Syntax

وهذه المفاهيم- على تعددها- يجمعها عدد من الخصائص:

- أنها تعبر عن "وحدات" للتحليل، لها دورها الحاسم في عملية التحليل اللغوي في مستوياته المختلفة، وكان الملجئ إليها إنما هو الحاجة إلى "وحدة " تحليلية، يمكن التعامل معها باعتبارها مقولة تصلح للتحليل والتركيب.
- ارتباطها الشديد بالدلالة.
- أنها وحدات تنتمي إلى اللغة الطبيعية[٢٧].
- أنها توظف في التحليل والتركيب.

على أنه يمكن القول إن هذه الوحدات لها تَحَقُّقُها الشكلي فيما عدا السيميم (الوحدة الدلالية)، وإن الديفيم المفترض ليس التحقق الفعلي للسيميم بل هو وحدة معجمية.

٥. **الجدوى:**

- **في التأليف:**
  - هي الوحدات الذرية التي ينبني منها التعريف في جملة دالة مؤدية لمحتوى الوحدات المعجمية.
  - الالتزام بقائمة محددة من الديفيمات وبقواعد تآلفها يؤدي إلى الالتزام بأنماط تعريفية ثابتة (هي بدورها محددة قبلا).
  - يؤدي الالتزام بأنماط تعريفية محددة إلى تحقيق الاتساق والبنية السليمة لمادة المعجم.
  - يمكّن من اشتقاق تعريفات المشتقات من تعريفات أصولها.
- **في التحليل:**

يمكن استخدامها في تطبيقات معالجة اللغات الطبيعية من خلال محلل التعريفات، حيث يكون التعامل مع كيانات Entities يمكن معالجتها آليا عن طريق محلل جُمَلي آلي Parser؛ استخلاصا وتصنيفا.

وتنتمي الوحدات التعريفية Defimes إلى إحدى المجموعات التالية:

المجموعة	**تحققاتها : أنماطها**
المجموعة الصرفية	سمات صرفية دلالية مقولات صرفية دلالية علاقات اشتقاقية دلالية

(٢٦) ويُقْصَد به الوحدة الدنيا التي تؤثر في تحديد معنى كلمة وتمييزه عن كلمة أخرى.

(٢٧) بخلاف تلك المفاهيم المشابهة التي لا تنتمي إلى اللغة الطبيعية، كوحدات اللغات الاصطناعية مثل لغات البرمجة، وكالديفيم الفرضية موضوع هذه الفقرة لانتمائها إلى اللغة الشارحة Meta Language.

المجموعة التركيبية	قيود انتقائية(فاعلية/ طبيعتها...الخ) بنية تركيبة للمحمولات بنية الحدث والبنية
المجموعة الدلالية	سمات دلالية / ذرات دلالية علاقات دلالية (أفقية – رأسية) مجال دلالي هل يمكن توظيف بنية الكواليا هنا
المجموعة الدالة على محددات التعريف	
الروابط hinges التي تربط المعرف بالتعريف	

يمكن تصور الديفيم باعتباره مفردات لغة شارحة، تعبر عن مدلولات وراء اللغة meta- language تستعمل في وصف البنية المعجمية الدلالية لمعجم. يضبط الديفيم اللغةَ ويضبط الأنماطَ بحيث تبقى هناك فراغات (غير ديفيمية، أو غير محددة، يمكن للمعجمي أن يملأها من معجم آخر). وعلى الاحتمال الثاني فإن ذلك يفترض وجود قائمتين معرفتين قبلا، هما: قائمة الديفيمات، وقائمة كلمات التعريف. ويكون التحليل عندئذ متوجها إليهما على الترتيب. وعلى ذلك يكون النمط التعريفي مرتبطا بمجموعة من الديفيمات. وعندئذ سيبرز التساؤل: وما جدوى التصور عندئذ ما لم يكن مستوعبا ؟

وهذه الفرضية يمكن أن تكون الرابط الوثيق بين محتوى التعريف ولغة التعريف، إذ هى الوحدة اللغوية التعريفية الحاملة لمحتوى تعريفي. على أن هذا المحتوى قد يكون تركيبيا أو دلاليا أو صرفيا.

٦. **الإجراءات:** (إجراءات التوصل إلى قائمة الوحدات التعريفية الديفيمات واستخلاصها).

التوصل إلى الوحدات التعريفية أهم نتائج التحليل للمدونتين المعتمدتين (مدونة النصوص، ومدونة التعريفات المعجمية)

- تحليل عينات تعريفية لاستخلاص الديفيمات المستخدمة من خلال استراتيجيات التعريف، وذلك باتخاذ مدونة من تعريفات أشيع معجمين عربيين حديثين هما المعجم الوسيط والمعجم العربي الأساسي.
- استخلاص الديفيمات من المقاربات الدلالية السابقة، وقد أفادت الدراسة من تصورات كاتز، وميلتشوك وبوسطيوفسكي في تمثيل المحتوى المعجمى الدلالي.
- التحليل الدلالي التركيبي للعينة، لاستخلاص الوحدات التي تؤدى بها المعلومات التركيبية الدلالية.
- التحليل الدلالي، لاستخلاص السمات الدلالية التي تحكم بنية المعجم والوحدات التي يمكن تمثيلها بها.
- التحليل الصرفي الدلالي: (المعجم أو الشبكة الصيغية)، وغايته استخلاص السمات الصرفية الدلالية للصيغ المزيدة، والمقولات الصرفية الاشتقاقية، إضافة إلى الوحدات التي تستخدم في اشتقاق تعريفات المشتقات من تعريفات أصولها.
- صياغة القائمة النهائية للديفيمات المستخلصة.
- صياغة قواعد التراكب والتآلف (وحدات واجبة، اقتراحات لوحدات ممكنة، احتمالات ممتنعة).
- اقتراح أنماط رئيسة وفرعية للتعريف.
- اقتراح آليات توليد التعريفات الفرعية من أصولها.

٧. **قواعد التآلف:** Defime Syntax

- للديفيمات قواعد تحكم تراكبها عند بناء التعريف، أو جزء التعريف، لتكوين نَوْعَي التعريف الرئيسي والفرعي[٢٨].
- توجد علاقات داخلية بين قائمة الديفيمات، لكونها تعد بمثابة معجم مصغر Sub Lexicon ، به من الكيانات (الدلالية والتركيبة) والعلاقات (الاشتقاقية والدلالية) ما بالمعجم العام.
- من هذه العلاقات- ولعله أشدها تأثيرا في العمل- التوارث Inheritance[٢٩]، ويعني الترابط الكلي لبنية المعجم، من خلال العلاقات التي تربط كياناته: مداخل ودلالات.
- قواعد التراكب والتآلف تبين(القيود الواجبة، الاقتراحات الممكنة، الاحتمالات الممتنعة).
- يمكن افتراض برنامج يكون بمثابة مدقق تركيبي Definition Checker للتعريفات أثناء عملية التحرير، ويكون جزءا من برنامج محرر التعريفات المقترح.

(٢٨) التعريف الرئيسي يكون للأصول، والفرعي يكون للمشتقات.

(٢٩) التوارث هنا مستخدم بمفهوم جيمس بوسطيوفسكي J. Pustejovsky بكونه البنية الكلية للمعجم والعلاقات التي تحقق هذا. J. Pustejovsky: Generative Lexicon.p:24

تصور عام للغة التعريف

قائمة الوحدات التعريفية (الديفيمات)	قواعد التآلف نحو التعريف	أنماط التعريفات (أبنية التعريف)		آليات توليد تعريفات المشتقات (المطردة)
		نمط رئيسي	نمط فرعي	
البنية المعجمية الدلالية				

٨. **الكفاية:**

مثل أي تصور ينبغي قياس درجة كفايته مقارنا بالتصورات المقترحة المنافسة (السابقة). والكفاية تعني مدى ما يحققه التصور النظري من دقة في تمثيل الظواهر اللغوية: ملاحظة ووصفا وتفسيرا. وقد تعددت معايير اختبار الفرضيات اللغوية لتغطي كل مجالات الظواهر المدروسة. فقد اقترح تشومسكي مستويات ثلاثة للكفاية في مجال النظرية النحوية[٣٠]، وأعاد جاكندوف توظيفها في مجال المعجمية
ونظرية المعجم[٣١]، وهي:

- الكفاية الملاحظية Observational Adequacy.
- الكفاية الوصفية Descriptive adequacy .
- الكفاية التفسيرية explanatory adequacy .

وقد تعرضت هذه الأفكار للمراجعة والتطوير. ومما ينبغي ذكره في هذا السياق إضافة بوسطيوفسكي[٣٢] مستوى آخر، هو:

- الكفاية الإمبيريقية (التجريبية) Empirical adequacy .

ويمكن أن تقترح معايير لقياس الكفاية اللسانية الحاسوبية لمنهجية تمثيل التعريفات، وهي:

- التحليلية: قابلية التعريفات للتحليل الآلي: Parsing .
- التركيبية: قابليتها التركيب.
- بساطة الوحدات التعريفية، ومخرجات التحليل، في مقابل التعقيد.
- الجدوى والإفادة في تطبيقات معالجة اللغات الطبيعية.
- واقعية التصور وقابليته للتطبيق.
- الاستيعاب للظاهرة اللغوية.

التصور والتطبيقات الحاسوبية

غاية التصور معالجة المحتوى المعجمي الدلالي بغرض تطبيق المخرجات في أنظمة معالجة اللغات الطبيعية وتطبيقاتها. فهذه الأنظمة والتطبيقات لها متطلبات في المحتوى، وأخرى في اللغة: فمشكلة معالجة اللغات الطبيعية مع المصادر المعجمية تنحصر في أمرين: المعارف اللغوية التي تحتاج إليها هذه المعالجة في تطبيقاتها، ولغة التمثيل المنضبطة التي يمكن التعامل معها حاسوبيا.

ولقد غدت القابلية للتوظيف في التطبيقات الحاسوبية معيارا من معايير الحكم على كفاية تصور نظري يهدف إلى تمثيل المحتوى المعجمي الدلالي.

التعريفات مصدرا للمعارف المعجمية الدلالية:

يحدد بوجريف Bran Boguraev وتيد بريسكو Ted Briscoe المعارف اللازمة لأنظمة معالجة اللغات الطبيعية N L P بأنها: المعارف النطقية (الفونولوجية) والصرفية والتركيبية والدلالية والذرائعية[٣٣]. وموضوع الدراسة يستبعد المعلومات النطقية لضآلة ارتباطه بالمحتوى المعجمي الدلالي، في حين يركز على الجوانب التركيبية والصرفية والدلالية. وبذا يكون التعريف المعجمي مصدرا لهذه الأصناف
من المعلومات اللغوية بالنسبة لأنظمة معالجة اللغات الطبيعية وتطبيقاتها.

جوانب فك اللبس في التصور المقترح:

من أهداف التصور المقترح تقديم المعارف اللغوية التي تلزم للمساعدة في فك اللبس[٣٤].

---

(٣٠) Chomsky, N. (1965) *Aspects of the Theory of Syntax*.

(٣١) Ray Jackendoff: Morphological and Semantic Regularities in the Lexicon p: 639.

(٣٢) J. Pustejovsky: Generative Lexicon p: 28.

(٣٣) Bran Boguraev & Ted Briscoe: Computational lexicography for natural language processing. P: 4:5

(٣٤) تجدر الإشارة هنا أن المقترح لا يدعي فك اللبس تماما، فاللبس جزء أساسي من سمات اللغة الطبيعية ولكن اقتراح إجراءات لغوية أو لسانية، وتقديم معلومات لسانية، كل ذلك يساعد في فك اللبس.

والمعلومات التي اهتم بها المقترح لتحقيق ذلك هي:

على المستوى الصرفي:

- فك اللبس المقولي (مثلا: اسم الفاعل، الصفة المشبهة، صيغة المبالغة، اسم المهنة[٣٥])
- فك اللبس الصيغي الممكن بين الكلمات ذات الصيغ المتشابهة.
- فصل المعاني عن طريق فصل الأسر الاشتقاقية على أسس من أصل الاشتقاق.
- المشتق المطرد والمتحول معجميا.

على المستوى التركيبي:

- بيان أنماط المقولة التركيبية.
- بيان أنماط البنية الحملية.
- توضيح قيود الانتقاء Selection restriction .
- توضيح أبنية الحدث Event structures.

على المستوى الدلالي:

- تعد محددات التعريف نوعا من توصيف الوحدة المعجمية، وتفريق الدلالات وفصل اشتباكاتها.
- التصاحبات مُكَمِّلا للتعريف: يمكن توظيفها في مجال تعلُّم الآلة Machine Learning.
- العلاقات الدلالية.
- السمات الدلالية والفروق اللغوية.
- اقتراح معايير لتمييز أنماط الوحدات المعجمية من حيث طبيعة محتواها الدلالي (العامة،الموسوعية، المصطلحية) واتخاذ ذا معيارا من معايير الاستقلال بمدخل.

في التحليل والتركيب: الفهم والتوليد:

وذلك من خلال مفهوم الوحدة التعريفية التي تعد وحدة للتحليل والتركيب.

يسهم المقترح- في حالة تطبيقها- في تسهيل تحويل المعارف المعجمية الدلالية إلى كيانات مفهومية Ontologies، وهو توجه ذو أهمية كبيرة في تطوير الويب الدلالي.

في المعجمية الحاسوبية:

- اقتراح محرر لتعريفات المعجم العربي يحتوي قواعد المنهجية المقترحة، ويحكم عمل محرري المعجم.
- اقتراح مولد آلي لتعريفات المشتقات، يشتق مادته من تعريفات الأصول.

لا يمكن إنجاز مولد التعريفات بدون إنجاز هذه الفرضية على المستوى النظري، ثم تطبيقها عمليا بحيث تتحدد قائمة منتهية Set final من الديفيمات وتنوعاتها الشكلية Allo-Defime بحيث يصبح بالإمكان عمل الدالة التي تكون أساس التوليد في المجال الحاسوبي.

الجديد في التصور المقترح

في تصور طبيعة المحتوى المعجمي الدلالي موضوع التمثيل في صورة التعريفات.

في لغة التعريف:

- اقتراح معجم وحدات التعريف وتصميم جوانبه وأسس بنائه.
- اقتراح نحو للتعريف يمكّن من التحليل والتركيب.
- اقتراح جوانب تحقق بنية المعجم الدلالية.

في الجانب الصرفي:

- افتراض نمطين من المشتقات: أصول وفروع، وإعادة بناء المدخل المعجمي العربي على أساس ذلك.
- اقتراح آليات لتوليد تعريفات المشتقات من تعريفات أصولها.
- توظيف المحتوى الصرفي في بناء المعجم.

على المستوى المعجمي:

- تعيير أنماط المداخل المعجمية.
- توسيع معايير فصل المشتركات اللفظية.

التوليد الآلي لتعريفات المشتقات:

يجد توليد المشتقات كثيرا من الطروحات الحاسوبية للحديث عنه: مفهومه وآلياته، وفي المقابل لا نجد حديثا عن توليد مواز لتعريفات المشتقات (المطردة)، وهو ما يعد ضرورة في اقتراح لأي مصدر معجمي موسع.

مشروع محرر التعريفات:

يدخل هذا المقترح ضمن اهتمام الصناعة المعجمية الحاسوبية Computational Lexicography بوصفه أداة لضبط لغة التعريف وتنميطه أثناء مرحلة التحرير.

(٣٥) نقصد باسم المهنة هنا: قسما من أسماء الفاعلين، وصيغ المبالغة تدل على حرفة أو مهنة أو وظيفة مثل (النجار، السقاء، الممرضة، الخ).

المراجع:

- **جوديث جرين:** علم اللغة النفسي، تشومسكي وعلم النفس. ترجمة وتعليق مصطفى التوني. المفهومية التي تجتمع فيها البنى الإدراكية للعالم.
- **وفاء كامل:** معجم التعابير الاصطلاحية العربية المعاصرة. أبو الهول للنشر- القاهرة ٢٠٠٧.
- **حسين نصار:** المعجم العربي نشأته وتطوره, ط ٤، ١٩٨٨ دار مصر للطباعة، القاهرة.
- **حلام الجيلالي:** تقنيات العريف في المعاجم العربية المعاصرة منشورات اتحاد الكتاب العرب، دمشق ، سوريا ١٩٩٩.
- **مجمع اللغة العربية بالقاهرة:**المعجم الوسيط ط٢, القاهرة, سنة ١٩٧٢. (إعداد لجنة من المجمع).

- Chomsky, N. (1965) Aspects of the Theory of Syntax.
- Kats & Fodor 1963 the structure of semantic theory.
- Dwight Bolinger: The Atomization of Meaning.

http://www.jstor.org/pss/411524

- Nida , Eugene, *A componential Analysis of Meaning. An introduction to semantic structures.*
- J. Pustejovsky: *Generative Lexicon.*
- Ray Jackendoff: *Morphological and Semantic Regularities in the Lexicon.*
- Bran Boguraev & Ted Briscoe: *Computational lexicography for natural language processing.*
- Susan Hunston: *Corpora in Applied Linguistics.*
- Ayto. J. R. (1983). *Semantic Analysis & Dictionary Definitions. In* R.R.K Hartmann (Ed.), *Lexicography: Principles & Practice.* Exeter.
- Bogreav, B. & Levin, B. (1993). *Models for Lexical knowledge. Bases.* In J. Pustejovsky (Ed.), *Semantics and the Lexicon.* Kluwer Academic Publishers
- Litkowski, K. C. (2005). *Computational Lexicons and Dictionaries.* In : *Encyclopedia of Language and Linguistics* (2nd ed.). Oxford, Elsevier Publishers.
- Svensén, B. (1993). *Practical lexicography :Principles and Methods of Dictionary-Making.* Translated from the Swedish. By Sykes, J. and Schofield, K. Oxford University Press

الدكتورة/ وفاء كامل

- أستاذة بكلية الآداب جامعة القاهرة.
- عضو مراسل بمجمع اللغة العربية بدمشق.
- خبيرة بمجمع اللغة العربية بالقاهرة.

- نشرت ٥٣ بحثا باللغتين العربية والانجليزية، وأشرفت على ٤٨ رسالة للماجستير والدكتوراه.
- حصلت على جائزة جامعة القاهرة التشجيعية عام ٢٠٠٤، وجائزة جامعة القاهرة التقديرية عام ٢٠١٣.
- ألّفت وترجمت ٩ كتب منها:(تراكب الأصوات في الفعل الثلاثي الصحيح)- (قصيدة الرثاء بين شعراء الاتجاه المحافظ ومدرسة الديوان: دراسة أسلوبية إحصائية)- (الباب الصرفي وصفات الأصوات)- (بحوث في العربية المعاصرة)- (المجامع العربية وقضايا اللغة)- (معجم التعابير الاصطلاحية في العربية المعاصرة)- (اتجاهات البحث اللساني)- (مدخل إلى اللغة).

الدكتور/ مُحسِن عبد الرَّازق علي رشوان

يشغل منصب أستاذبقسم الإلكترونيات والاتصالات الكهربائية في كُلِّيَّة الهندسة–جامعة القاهرة. تخرَّجَفي هذا القِسمِ عام ١٩٧٧ وكان الأول على دفعته، وحصلَ على درجة الماجستير الأولى من القِسم ذاتِه عام ١٩٨٠. حصل على درجة الماجستير الثَّانية في الهندسة الكهربائية من جامعة كارلتون بكندا، ثم حصلَ على درجة الدُّكتوراه من جامعة كوين بكندا عام ١٩٨٧؛ كما حصلَ على درجة الماجستير في إدارة الأعمال عام ٢٠٠٢ من الأكاديمية العَرَبِيَّة للعُلُوم والتكنولوجيا والنَّقل البَحَرِيّ في مِصر.أشرف على أكثر من خمسٍ وثمانينأطروحة عِلميَّة(ماجستير ودكتوراه)، ونشر ما يربُو علىمائة بحثًا في دوريَّاتٍ علميَّةٍ مُتَخَصِّصة ومؤتمراتٍ دوليَّة؛ كما سَجَّلَ أربع براءات اختراع.يدير الشَّركة الهندسيَّة لتطوير النُّظُم الرَّقمِيَّةRDI المُتَخَصِّصة في مجال تقنيات اللُّغة العربيَّة، وحصلَ من خلالها على العديد من الجوائز

المحليَّة والدَّوليَّة، ومنها: الجائزة العالميَّة *WSA 2007* عن تقنية حفص لتعليم أحكام تلاوة القُرآن الكريم. ساهم مع آخرين لانشاء مؤسسة لا تهدف للربح لدعم تقنيات اللغة العربية *ALTEC (www.ALTEC-Center.org)*