ترميز الظواهر الدلالية في المعجم الحاسوبي العربي

يوسف أبو عامر* ، وفاء كامل**
*كلية الآداب، جامعة القاهرة، مصر
ysfaboamer@cu.edu.eg
**كلية الآداب، جامعة القاهرة، مصر
wafkamel@yahoo.com

المستخلص - الدلالة –في أبسط تعريفاتها- "دراسة المعنى"، وهي أقل مستويات اللغة العربية خضوعًا للبحث في مجال اللسانيات الحاسوبية؛ إذ تنطوي على جملة من الظواهر التي تخلق قدرًا كبيرًا من التعقيد كالترادف Synonymy، والتضاد Antonym، والاشتمال Hyponymy، وغيرها، كما نتج عن الثراء الصرفي والغموض الكتابي بسبب غياب التشكيل في أكثر الأحيان نسبة كبيرة من المتجانسات اللفظية، ومن ثم مزيد من الغموض في المعنى.

تحتاج بعض نظم الفهم الآلي للغات الطبيعية -كاسترجاع المعلومات Information Retrieval، 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 يمكنها من تحليل "معنى" الجملة.

تتناول هذه الدراسة طبيعة المحتوى الدلالي الذي يُرَمَّز في المعاجم المصممة لتكنولوجيا اللغات الطبيعية -بشكل عام- تبعًا لتوصيات المجموعة الاستشارية الخبيرة في معايير هندسة اللغة، وتهدف بشكل أساسي إلى تطبيق المنهجية المقترحة لتمثيل هذه الظواهر الدلالية على مداخل المعجم الحاسوبي العربي المصمم لتطبيقات المعالجة الآلية.

وقد قسمت الظواهر الدلالية –بحسب أهميتها لتطبيقات المعالجة الآلية- إلى ثلاثة مستويات: عالية، ومتوسطة، ومنخفضة. تركز الدراسة على الظواهر ذات الأولوية الأعلى؛ لأنها تُستخدم كثيرًا في تطبيقات المعالجة الآلية.

الكلمات المفاتيح
المعجم الحاسوبي – المعالجة الآلية للغات الطبيعية – الظواهر الدلالية – ترميز

1 المقدمة

يمثل المعجم ركيزة أساسية يعتمد عليها بناء النظم المختلفة للمعالجة الآلية للغات الطبيعية؛ فالأداء النهائي لنظم المعالجة الآلية يعتمد على كفاءة المعجم المرتبط بها؛ إذ تؤدي المعلومات الصرفية، والنحوية، والدلالية -التي يُزَوِّد المعجمُ النظامَ بها- دورًا مهمًّا في هذا الأداء، ومع ذلك قد تتنوع مستويات التفصيل، ونوعية المعلومات التي يحتوي عليها المعجم باختلاف نظام المعالجة.

ظلت الظواهر الصرفية النحوية مركز اهتمام الباحثين في العمل المعجمي العربي الحاسوبي، ولم تحظ الدلالة –على أهميتها- باهتمام الباحثين، وهو أمر عائد إلى درجة التعقيد العالية التي تتسم بها منظومة الدلالة التي تنطوي على جملة من الظواهر التي تخلق قدرًا كبيرًا من التعقيد كالترادف، والتضاد، والاشتمال، وغيرها.

تتناول هذه الدراسة المعلومات الدلالية الذي تُرَمَّز في المعاجم المصممة لتكنولوجيا اللغات الطبيعية -بشكل عام- وفقًا لتوصيات المجموعة الاستشارية الخبيرة في معايير هندسة اللغة، وتهدف بشكل أساسي إلى تطبيق المنهجية التي تقترحها المجموعة لتمثيل هذه الظواهر الدلالية على مداخل المعجم الحاسوبي العربي المصمم لتطبيقات المعالجة الآلية.

تنقسم هذه الدراسة إلى أربعة أجزاء: يتناول الأول منها أهمية الدلالة بشكل عام للإنسان وللآلة، في حين يعرض الثاني تقسيم المجموعة الاستشارية للمعلومات الدلالية تبعًا لأهميتها لتطبيقات المعالجة المختلفة، ويتناول الجزء الثالث مجموعة الظواهر الدلالية التي تحتاجها معظم نظم المعالجة الآلية وكيفية تمثيلها في المعجم الحاسوبي العربي، وأخيرًا تعرض الخاتمة النتائج التي توصلت إليها الدراسة.

2 أهمية الدلالة للإنسان وللآلة

الدلالة –في أبسط تعريفاتها- "دراسة المعنى"، وقد اهتمت الدراسات الدلالية بالمعنى على مستوى الكلمات والمورفيمات، وطبيعة العلاقات التي تنشأ بينها كالاشتمال، والترادف، والتطابق، وغيرها في إطار ما يعرف بالدلالة المعجمية(*) Lexical semantics. كما اهتمت بدراسة المعنى، والعلاقات الدلالية التي تنشأ بين الوحدات اللغوية التي تتجاوز مستوى الكلمة إلى العبارات والجمل؛ "إذ يقرر مبدأ التأليف The principle of compositionality أن معنى العبارة أو الجملة يعتمد على معاني كلماتها، والعلاقات النحوية بينها[8].

إن المعلومات المتعلقة بمعاني الكلمات تُخْتَزن في المعجم العقلي لدى البشر، تمامًا كما تُخْتَزن المعلومات الصرفية والنحوية. ففي جملة "كتب محمد الدرس"، إذا كان الفعل "كتب" موجودًا في المعجم العقلي، يدرك الإنسان أن شخصًا ما كتب شيئًا؛ إذ تخبره معرفته بمعنى الفعل "كتب" أن حيوانًا لا يمكنه الكتابة. وتتضمن المعرفة بالفعل أيضًا أن الشخص المشار إليه كائن حي، وهو الكاتب؛ أي الفاعل، وأن شيئًا وقع عليه فعل الكتابة. هذه المعلومات تُمَثَّل بعض الخصائص الدلالية Semantic properties للكلمة على النحو الذي يصطلح عليه متحدثو اللغة.

(*) يشيع استخدام مصطلح الدلالة المعجمية Lexical semantic في مقابل الدلالة الشكلية للإشارة إلى دراسة كلمات المحتوى Content words، التي لها محتوى وصفي يحدد أنواع الكيانات والأحداث التي تدل عليها، للمزيد انظر:
Kearns, K. Lexical Semantics in Aarts, B. & Mcmahon, A. (2006). The Handbook of English Linguistics. Blackwell Publishing Ltd.

فالمعجم –إذن- هو الجزء الذي يحتوي على المعرفة التي يمتلكها المتحدثون عن الكلمات المفردة والمورفيمات، بما في ذلك السمات الدلالية؛ فالكلمات التي تشترك في سمة دلالية ما يقال إن لها "التصنيف الدلالي Semantic class" نفسه، كالتصنيف الدلالي للكلمات "المؤنثة". وفي بعض الأحيان يمكن الاستدلال على سمة دلالية بوجود سمة دلالية أخرى أو غيابها، فالكلمات التي لها التصنيف الدلالي "بشري" لها التصنيف الدلالي "حي"، وتفتقر –في الوقت ذاته- إلى التصنيف الدلالي "حيوان"[8].

وكما يحتاج الإنسان إلى معرفة بمعاني الكلمات، وطبيعة العلاقة الدلالية بين العناصر المكونة للجملة، على نحو يسمح له بمعالجتها، تتطلب -أيضًا- حوسبة اللغة ولوجًا إلى معاني الكلمات، باعتبارها جزءًا من المعرفة اللغوية التي تساعد على التعامل مع الكلمات بشكل يقترب من التعامل البشري معها. وإذا كان المحتوى الصرفي كفيلًا بأن يوضح للحاسوب أن كلمة "كتب" تشير إلى نوع معجمي معين هو "الفعل"، وإلى حدث وقع في زمن مضى وانتهى، وأن مَن قام به مفرد مذكر، فإن الحاسوب يظل مفتقرًا إلى معرفة بمعنى الفعل، وهو "ما يمكن تحديده من خلال هذه الخصائص؛" إذ "يمكن تحديد معاني الأسماء، والأفعال، والصفات، والظروف –كلمات المحتوى Content words- حتى بعض الكلمات الوظيفية من خلال هذه الخصائص الدلالية[8].

3 تقسيم المجموعة الاستشارية للمعلومات الدلالية تبعًا لأهميتها

قدمت المجموعة الاستشارية الخبيرة في معايير هندسة اللغة (EAGLES) اقتراحًا بمجموعة من المعايير للترميز الدلالي في المعاجم والمدونات النصية[7]. وقد اشتملت هذه المعايير على المعلومات المفاهيمية الأساسية، كما قسمتها –بحسب أولوية ترميزها- ثلاثة مستويات، اعتمادًا على أهميتها للتطبيقات المختلفة[7]:

أ. *المستوى الأول: أولوية عالية High priority*

يشتمل هذا المستوى على المعلومات التي تُستخدم في معظم التطبيقات.

ب. *المستوى الثاني: أولوية متوسطة Medium priority*

يشتمل هذا المستوى على المعلومات التي تُستخدم تجريبيًّا في بعض التطبيقات.

المستوى الثالث: أولوية منخفضة Low priority

يشتمل هذا المستوى على المعلومات غير المستخدمة في الوقت الراهن، ولكنها ذات قيمة محتملة في تطوير بعض التطبيقات.

ويحتوي الجدول التالي رقم (1) على تقسيم المجموعة الاستشارية للمعلومات الدلالية تبعًا لأهميتها لتطبيقات المعالجة المختلفة.

الجدول رقم (1)
تقسيم المجموعة الاستشارية للمعلومات الدلالية تبعًا لأهميتها لتطبيقات المعالجة المختلفة

	أنواع الكلمة الأساسية	الاشتمال والترادف	التضاد	الجزئية	الإطار الدلالي	زمن الفعل	العلاقات الزمنية	الجهة	مجال الموضوع	علاقات الاقتران	الكمية
الترجمة الآلية	*	*		*	*	*	*	*	*	*	*
استرجاع المعلومات	*	*		*	*						
استخلاص المعلومات	*	*		*	*	*		*			
التلخيص الآلي	*	*							*		
إنتاج اللغة	*	*	*	*	*	*	*	*	*	*	*
تعرف التجمع اللفظي	*	*			*				*		
إزالة لبس المعنى	*	*	*		*				*	*	
أسماء الأعلام	*	*		*							
الإعراب	*	*			*			*		*	
الإحالة المشتركة	*	*			*						*

الأولوية العالية
أنواع الكلمة الأساسية Base word types (BWT)
الاشتمال والترادف Hyponymy/ Synonymy
الإطارات الدلالية Semantic frames

الأولوية المتوسطة
الجزئية Meronymy
التضاد Antonymy
مجال الموضوع Subject domain
علاقات الاقتران Co-occurrence relations
زمن الفعل Tense
العلاقات الزمنية Time (Temporal relations)
الجهة Aspect
الكمية Quantification

4 الظواهر الدلالية وتمثيلها في المعجم الحاسوبي العربي

تتناول الدراسة في هذا الجزء المعلومات الدلالية ذات الأولوية العالية، تلك التي تستخدم في معظم تطبيقات المعالجة الآلية، وهي: أنواع الكلمة الأساسية، والترادف، والاشتمال، والإطار الدلالي، كما تتناول أهم الظواهر ذات الأولوية المتوسطة، تلك التي تُستخدم تجريبيًّا في بعض التطبيقات، ومنها: الجزئية، والتضاد، ومجال الموضوع، وعلاقات الاقتران.

أ. *أنواع الكلمة الأساسية Base Word Types (BWT)*

في شبكة المفردات الأوروبية EuroWordNet تم عمل قائمة معيارية لهذه المفاهيم الرئيسية، وأصبح لها دور مهم في شبكات مفردات اللغات: الإنجليزية، والفرنسية، والألمانية، والإيطالية، والهولندية، والتشيكية، والإستونية، ومعجم اللغة السويدية. وتمثل هذه القائمة معظم الأنواع الأساسية، والعناصر الدلالية في كل لغةٍ، ولذلك أطلق عليها "المفاهيم الأساسية لشبكة المفردات الأوروبية."[7] وقد أجرت المجموعة بعض التعديلات البسيطة على هذه القائمة؛ ليصبح الناتج النهائي مجموعة مؤلفة من أربعة وسبعين (74) مفهومًا موزعًا على عدد من المستويات: ينقسم المستوى الأول منها إلى: "**كيان Entity**" ويضم الأشياء الملموسة، و"**مفهوم Concept**" ويضم المفاهيم والأفكار، و"**حدث Event**" وهي الأحداث التي تنطوي على تغير، و"**حالة State**" وهي المواقف الثابتة.

ويتفرع كل نوع من هذه الأنواع إلى مستويات أعمق، قد تصل في بعض الأحيان إلى خمسة مستويات، على النحو التالي:

1 . كيان (ENTITY)
1 . 1. أصلي (ORIGIN)
1 . 1 . 1 . طبيعي (NATURAL)
1 . 1 . 1 . 1 . حي (ANIMATE)
1 . 1 . 1 . 1 . 1 . إنسان (HUAMN)
1 . 1 . 1 . 1 . 1 . 2 . حيوان (ANIMAL)
1 . 1 . 1 . 1 . 1 . 3 . نبات (PLANT)
1 . 1 . 1 . 1 . 2 . غير حي (INANIMATE)
1 . 1 . 1 . 2 . صناعي (ARTIFACT)

1 . 2. توليف (COMPOSITION)
1 . 2 . 1 . كل (WHOLE)
1 . 2 . 2 . مجموعة (GROUP)
1 . 2 . 3 . جزء (PART)

الشكل رقم (1)
جزء من تقسيم الأنواع الأساسية للكلمة

وتعد الأنواع الأساسية للكلمة من أهم المفاهيم الدلالية –إن لم تكن أهمها على الإطلاق- لتطبيقات معالجة اللغة آليًّا؛ إذ تُستخدم في معظم هذه التطبيقات، بدءًا من الترجمة الآلية، ومرورًا باسترجاع المعلومات، واستخلاصها، وانتهاءً بالتلخيص الآلي للنصوص، وكذلك في التكنولوجيا المساعِدة كإزالة اللبس، وإنتاج اللغة، والإعراب، وغيرها.

ب. *الاشتمال والترادف Hyponymy/ Synonymy*

تُمَثِّل علاقتا الاشتمال والترادف جوهر المعاجم الدلالية. ويقصد بالاشتمال العلاقة التي تتحقق بين وحدات محددة وأخرى عامة بحيث تكون الأولى مُشْتَمَلةً في الثانية؛ أي أن الأولى نوع من الثانية، فالقطة نوع من الحيوان، وكذلك الحال للمزمار في ضوء علاقته بالآلة، والكرسي بالأثاث[6].

وتبعًا لكروز تتحق علاقة الاشتمال بين الاسمين (أ) و(ب)، التي يكون فيها (أ) نوعًا من (ب) إذا استلزمت هذه العلاقة –من جانب واحد فقط- أن (أ) نوع من (ب)، وليس العكس، كما في الأمثلة التالية[5]:

هذا كلب	تستلزم من جانب واحد	هذا حيوان
هذا فحل	تستلزم من جانب واحد	هذا حصان
هذه وردة قرمزية	تستلزم من جانب واحد	هذه وردة حمراء

ويتيح التعدي في الاشتمال إمكانية تنظيم معاني كلمات اللغة في شكل شبكة هرمية، أو مخطط تصنيفي. ويوضح الشكل (2) أن الشبكات الهرمية تمثل بِنى فعالة؛ إذ يمكن عمل بنية توارث Inheritance structure لعدد كبير من معاني الكلمات التي ترتبط بشكل مباشر أو غير مباشر بالمستويات العليا.

أما مصطلح الترادف، فيشير إلى نوع رئيسي من علاقات المعنى بين الوحدات المعجمية؛ فالوحدات المعجمية التي لها المعنى نفسه هي وحدات مترادفة Synonyms. ولا يعني الترادف بين وحدتين معجميتين أن ثمة تطابقًا كاملًا بينهما في المعنى، بل يُستخدم مصطلح الترادف إذا كان هناك تقارب في المعنى بين وحدتين معجميتين بحيث يسمح هذا التقارب باختيارٍ بينهما في بعض السياقات، دون تغيير في المعنى الكلي للجملة[6].

Entity (كيان)
……. Animate (حي)
……. …….. Human (إنسان)
……. …….. Plant (نبات)
……. …….. Animal (حيوان)
……. Inanimate (غير حي)
……. …….. Natural (طبيعي)
……. …….. Artifact (صناعي)
……. …….. ……. Building (مبنى)
……. …….. ……. ……. Apartment (شقة)
……. …….. ……. ……. Bakery (مخبز)
……. …….. ……. ……. Church (كنيسة)
……. …….. ……. ……. Hotel (فندق)
……. …….. ……. ……. House (منزل)
……. …….. ……. ……. Mosque (مسجد)
……. …….. ……. ……. School (مدرسة)
……. …….. ……. ……. Theater (مسرح)
……. …….. ……. ……. Temple (معبد)
……. …….. ……. ……. University (جامعة)
……. …….. ……. ……. etc

الشكل رقم (2)
عمل بنية توارث لعدد من معاني الكلمات التي ترتبط بشكل مباشر أو غير مباشر بالمستويات العليا

ويتم ترميز المعلومات الدلالية المتعلقة بالأنواع الأساسية، وطبيعة العلاقة المعجمية بين هذه الأنواع، وكذلك المرادفات والمُشتَمَلات، على النحو الموضح بالمثالين التاليين:

[ENTRY: سيارة
BASE-TYPE-INFO: [BASE-TYPE: ARTIFACT
LX-REALTION: LX-HYPONYM]
[BASE-TYPE: OBJECT
LX-REALTION: LX-HYPONYM]
[BASE-TYPE: WHOLE
LX-REALTION: LX-HYPONYM]]
[BASE-TYPE: INSTUMENT
LX-REALTION: LX-HYPONYM]]
SYNONYMS: [عربة، مركبة، حافلة، شاحنة، أتوموبيل]
HYPONYMS:[سيارة إسعاف، مقطورة، ميكروباص، سيارة نقل ثقيل، سيارة أجرة، سيارة نقل عام، سيارة ملاكي، سيارة رياضية]
[ENTRY: حمار
BASE-TYPE-INFO: [BASE-TYPE: NATURAL
LX-REALTION: LX-HYPONYM]
[BASE-TYPE: ANIMATE
LX-REALTION: LX-HYPONYM]
[BASE-TYPE: ANIMAL
LX-REALTION: LX-HYPONYM]]
SYNONYMS: بغل
HYPONYMS: أتان، جحش، حمار وحشي

الشكل رقم (3)
نموذجان لترميز الأنواع الأساسية والمترادفات والمشتملات للمدخلين: سيارة، وحمار.

تُستخدم علاقتا الاشتمال والترادف في العديد من تطبيقات المعالجة الآلية التي تتعامل مع المحتوي؛ إذ يمكن من خلال هاتين العلاقتين:
- الوصول إلى التنوع المعجمي عن طريق إعطاء كلمات بديلة أو قريبة يمكن استخدامها للتعبير عن المعنى نفسه، أو معنى قريب.
- عمل استنتاجات دلالية بتوليد الخصائص الدلالية من المُشتَمَلات[7].

ويمكن الإفادة من ذلك في استرجاع المعلومات واستخلاصها، والترجمة الآلية، والتلخيص الآلي، وإنتاج اللغة، ولذلك يمثل ترميز علاقتي الترادف والاشتمال أولوية عالية في توصيات المجموعة الخبيرة[7].

ت. الجزئية Meronymy

يستخدم مصطلح الجزئية في علم الدلالة جزءًا من دراسة علاقات المعنى التي تربط الوحدات المعجمية، ويقصد به العلاقة بين الأجزاء والكليات، التي يكون فيها (أ) جزءًا من (ب). وتختلف هذه العلاقة -بشكل خاص- مع علاقة "الاشتمال" التي يكون فيها (أ) نوعًا من (ب)[6].

وتمثل أجزاء الجسم، وأجزاء السيارة أشهر الأمثلة التي توضح علاقة الجزئية، فالإصبع جزء من اليد، واليد بدورها جزء من الذراع، والذراع جزء من الجسد. ويطلق على العلاقة المقابلة اسم "الكلية Holonymy"، وفيها يكون الجسد كلًّا للذراع، ويكون الذراع بدوره كلًّا لليد، واليد كذلك للإصبع. وتجدر الإشارة إلى أن علاقة الكلية تمثل أولوية متوسطة تبعًا لتوصيات المجموعة الخبيرة، وترتبط أهميتها بتطبيقات الترجمة الآلية، واسترجاع المعلومات، واستخلاصها، وإنتاج اللغة. وقد اقْتُرح ترميزها على النحو التالي:

[ENTRY:
MERONYM: [MER-TYPE: (member | substance | part)
HOLS: hols]
HOLONYM: [HOL-TYPE: (member | substance | part)
MER: mers]]

ويمكن تطبيق ذلك على المدخل المعجمي "رأس"

[ENTRY: رأس
MERONYM: [MER-TYPE: (part)
HOLS:جسد]

الشكل رقم (4)
نموذج لترميز علاقة الجزئية للمدخل "رأس"

ث. التضاد Antonym

يشير مصطلح التضاد إلى كل أنواع التعارض الدلالي، سواء بين المتقابلات المتدرجة Graded antonyms، مثل: صغير وكبير، وطويل وقصير، وساخن وبارد؛ إذ توجد درجات من الاختلاف، والمتقابلات غير المتدرجة Ungraded antonyms، مثل: حي وميت، وأعزب ومتزوج، وحاضر وغائب[6]. وتمثل علاقة التطابق هي الأخرى أولوية متوسطة تبعًا لتوصيات المجموعة الخبيرة، وترتبط أهميتها بتطبيقات مثل إنتاج اللغة، وإزالة اللبس.

[ENTRY: طويل
ANTONYM: قصير

ج. مجال الموضوع Subject Domain

اعتمدت المجموعة الخبيرة ترميز قاموس كمبريدج الدولي للغة الإنجليزية CIDE مرجعًا فيما يتعلق بترميز مجال الموضوع Subject domain، واختارت الترميزات التي تَرِد كثيرًا في عدد من القواميس(*) كما يظهر من الأمثلة الموضحة بالشكل رقم (5).

Building & Civil Engineering (البناء والهندسة)
Building (البناء)
Building Tools (أدوات البناء)
Interior Decoration (هندسة الديكور)
Buildings (المباني)
Archeology (علم الآثار)
Architecture (هندسة المعمار)
Media & Publishing (الإعلام والنشر)
Media (الإعلام)

الشكل رقم (5)
بعض الأمثلة من الحقول الدلالية المُرمّزة في قاموس كمبريدج الدولي للغة الإنجليزية

ح. علاقات اقتران الكلمة

تَسْتَخدم المجموعة الخبيرة هذا المصطلح للإشارة إلى أية علاقة تجمع بين وحدتين معجميتين تردان في الوقت ذاته في تركيبٍ واحدٍ، مثل علاقة الفعل والمكمل، والعلاقات التي تجمع بين مُكوِّنات المركبات، وبذلك يغطي المصطلح مجموعةً متنوعة من التراكيب اللغوية، كالتصاحبات اللفظية Collocations، والكلمات المركبة، والقيود الانتقائية[7]. وفي هذا الجزء، تتناول الدراسة تمثيلًا لعلاقات الاقتران باثنتين من أهم هذه العلاقات بالنسبة لتطبيقات المعالجة الآلية وهما: التصاحب اللفظي، والتعبير الاصطلاحي، في حين تتناول موضوع "القيود الانتقائية"، وطريقة ترميزه في المعجم باعتباره عنصرًا من عناصر "إطار المسند Predicate frame". وهو ما فعلته كذلك المجموعة الاستشارية الخبيرة في توصياتها.

1) التصاحب اللفظي

التصاحب اللفظي هو تجمع من الكلمات المتجاورة نحويًّا، والمترابطة في الاستعمال، يشكل مركبًا منسجمًا من الناحية الدلالية. ويكون معنى التصاحب اللفظي مرتبطًا بمجموع معاني الكلمات المكوِّنة له، ومن أمثلته: السوق الحرة، والسيولة المرورية، والرسوب الوظيفي، والبطالة المُقنَّعة[1].

2) التعبير الاصطلاحي

مجموعة ثابتة من الكلمات، تشكل تعبيرًا يحمل معنى خاصًّا لا يمكن استنتاجه من مجموع معاني كلماته المفردة، مثل: حجر الزاوية، ودمه أزرق، وزوبعة في فنجان[1]. ولعل أبرز المعايير التي يُحدَّد على أساسها التعبير الاصطلاحي:

- **ألا يكون معنى التعبير تحليليًّا Non-compositionality**: أي لا تكون مكوناته محتفظة بمعناها المعجمي داخل التعبير، كما لا تحتفظ بدلالتها المجازية المعتادة منها[1]، وعلى هذا لا يمكن التنبؤ بمعنى التعبير الاصطلاحي من خلال عناصره.
- **عدم إمكانية استبدال عناصره Non-substitutability**: فلا يمكن استبدال أحد عناصر التعبير الاصطلاحي بعنصر آخر مرادف له.
- **عدم القابلية للتعديل Non-modifiability**: لا يقبل كثير من التصاحبات اللفظية التعديل سواء بإضافة مواد معجمية، أو عن طريق التحويلات النحوية[12].

(*) من هذه القواميس:

- Longman Dictionary of Contemporary English (LDOCE).
- Cambridge International Dictionary of English (CIDE).
- Chambers 20th Century Dictionary.

وتظهر أهمية ترميز التصاحبات اللفظية، والتعابير الاصطلاحية في العديد من التطبيقات، منها إنتاج اللغة؛ إذ يساعد تعرُّفها على التأكد من صحة المُخرجات، وعلى تجنب الأخطاء. والأمر ذاته ينطبق على الترجمة الآلية، إذ يسهم في دقة الترجمة من اللغة المصدر إلى اللغة الهدف. وقد اقترحت المجموعة الخبيرة ترميز هذا النوع من العلاقات على النحو الموضح بالشكل (6).

```
Collocation →
   [COLLOCATE: word-sense-id
              DIRECTION: (left | right)
              WORD-DISTANCE: [LOWER-LIMIT: integer
                                        UPPER-LIMIT : integer]
              DEPENDENCY: (h2d | d2h | d2d | h2h)
DEP-TYPE: dependency-type
PRPBABILITY: probability-value
```

الشكل رقم (6)
الترميز المقترح لعلاقات اقتران الوحدات المعجمية تبعًا لتوصيات المجموعة الخبيرة

وقد تتعدد مصاحبات المدخل المعجمي، ومن ثم تكون لكل مصاحب قائمة تضم أزاوجًا من الخصائص -على النحو الموضح بالشكل السابق- التي تحدد شكل التصاحب وطبيعته، وبيانها كالتالي:

- تعبر الخاصية "المصاحب COLLOCATE" عن الكلمة المصاحبة للمدخل المعجمي.
- تعبر الخاصية "الاتجاه DIRECTION" عن موقع المصاحب يمينًا أو يسارًا للمدخل المعجمي.
- تعبر الخاصية "مسافة الكلمة WORD-DISTANCE" عن المسافة بين المدخل والمصاحب عن طريق خاصيتين تحددان الحد الأدنى LOWER-LIMIT، والحد الأقصى UPPER-LIMIT.
- تصف الخاصية "الاعتماد DEPENDENCY" أشكال الاعتماد بين المدخل المعجمي والمصاحب. وقد اعتمدت المجموعة أربعة أشكال، اعتمادًا على التفصيل المعجمي لمشروع "SPARKLE"[*]، هي:
  - من الرأس إلى التابع (h2d)؛ أي مصدر الاختيار هو الرأس، وهدف الاختيار معتمد عليه نحويًّا.
  - من التابع إلى الرأس (d2h)؛ أي يعتمد مصدر الاختيار نحويًّا على هدف الاختيار.
  - من التابع إلى التابع (d2d)؛ أي يعتمد مصدر الاختيار، وهدف الاختيار كلاهما على عنصر آخر.
  - من الرأس (h2h)؛ أي مصدر الاختيار، وهدف الاختيار كلاهما رأسان لتركيبَي اقترانٍ مختلفين، ولكنهما متطابقان.
- تعبر الخاصية "نوع الاعتماد DEP-TYPE" عن الوظيفة النحوية باعتبار الرأس. وتُحَدِّدُ العلاقاتُ النحويةُ الاعتمادَ النحويَّ الذي يربط بين الرأس والتابع، وتنتظم هذه العلاقات في شكل هرميٍّ على النحو الموضح بالشكل رقم (7).
- تعبر الخاصية "نوع الاعتماد DEP-TYPE" عن الوظيفة النحوية باعتبار الرأس.

وتُحَدِّدُ العلاقاتُ النحويةُ الاعتمادَ النحويَّ الذي يربط بين الرأس والتابع، وتنتظم هذه العلاقات في شكل هرميٍّ على النحو الموضح بالشكل رقم (7).

```
dependency-type → arg | mod | arg_mod
arg → subj | comp
subj → ncsubj | xsubj | csubj
comp → obj | clausal
obj → dobj | obj2 | iboj
clausal → xcomp | ccomp
mod → ncomd | xmod | cmod
```

الشكل رقم (7)
أنواع الاعتماد بين الرأس والتابع في شكل هرميٍّ

- تعبر الخاصية "الاحتمال PROBABILITY" عن قوة الارتباط بين الكلمتين في واحدة أو أكثر من المدونات.

ويمكن التمثيل لذلك بالتعبير الاصطلاحي "حجر الزاوية"، مع الأخذ في الاعتبار إمكانية أن تكون إحدى المفردتين هي المدخل المعجمي، ومن ثم تتصاحب معها الكلمة الأخرى، وسنفترض هنا أن كلمة "حجر" هي المدخل:

```
Entry: حجر
   [COLLOCATE: الزاوية-1
              DIRECTION: left
              WORD-DISTANCE: LOWER-LIMIT: 0
                                        UPPER-LIMIT : 1
              DEPENDENCY:d2h
DEP-TYPE: mod]
```

الشكل رقم (8)
التمثيل المعجمي للتعبير الاصطلاحي "حجر الزاوية"

([*]) تمثل مجموعة الحروف اختصارًا (Shallow Parsing and Knowledge Extraction for Language Engineering)، وهو مشروع بحثي حول تطور المناهج والتقنيات الخاصة بالتطبيقات المعتمدة على المصادر المعجمية المأخوذة (شبه) آليًّا من النصوص.

خ. الإطار الدلالي Semantic Frame

اعتمدت المجموعة الخبيرة عددًا من العناصر تُشَكِّل الإطارَ الدلالي، هي: بنية المحمول، وإطارات التصنيف الفرعي، والقيود الانتقائية– باعتبارها امتدادًا طبيعيًّا لإطارات التصنيف الفرعي- والوظائف الدلالية Thematic roles، والقالب الدلالي Semantic templates، والتصنيفات الدلالية للفعل Verb semantic classification.

وتقتصر الدراسة في هذا الجزء على إطارات التصنيف الفرعي، والوظائف الدلالية، والقيود الانتقائية؛ باعتبارها العناصر الأهم –على مستوى الترميز- بين عناصر إطار المسند.

1) إطارات التصنيف الفرعي Subcategorization Frames

تُعَبِّر الأفعال عن محمولاتها الدلالية Semantic arguments بطرق نحوية مختلفة؛ فمجموعة معينة من الأبواب النحوية التي يمكن أن تظهر مع الفعل تُسمَّى "إطارَ تصنيفٍ فرعي". ويعد التصنيف الفرعي للفعل (المسند) عنصرًا مفتاحيًّا للمدخل المعجمي؛ لأن معظم النظريات النحوية الحديثة –إن لم تكن جميعها- تُخَطط للبنية النحوية من المعجم، وبذلك يتعين على المعرب -الذي يفيد من النحو المعجمي- أن يَلِج إلى معجمٍ شاملٍ يُرَمِّز عدد محمولات الفعل وأنواعها[3].

وتظهر أهمية هذا الترميز بصورة واضحة في الإعراب؛ إذ يَعزو كارول خمسين بالمائة (50%) من أخطاء الإعراب إلى غياب إطارات التصنيف الفرعي[4]. وإذا كان المعرب يستخدم التقنيات الإحصائية لترتيب التحليلات، فمن الضروري كذلك أن يرتب المعجم معدلات التكرار المرتبطة بأنواع التصنيفات المختلفة لكل فعل، وهذا بالإضافة إلى فائدته في كل تطبيقات المعالجة الآلية التي يمكن أن تفيد من المعلومات المتعلقة ببنية المسند والمحمول Predicate-argument structure، كالترجمة الآلية، واستخلاص المعلومات. وقد صمم عدد من معاجم التصنيف الفرعي للغة الإنجليزية، مثل معجم كومبلكس Complex الذي طوره جريشمان وآخرون، ومعجم ANLT، الذي طوره بوجوريف وبريسكو[3]. كما قدم عطية وآخرون عملًا مماثلًا للغة العربية[2].

ونستطيع أن نتبيَّن أهمية هذه التصنيفات لمهام الإعراب -مثلًا- إذا نظرنا إلى إعراب المثالين التاليين:

أخبرت زوجها أن الطفل في المكان المتفق عليه.

وجدت الزوجة المكان المتفق عليه.

فإذا علمنا أن للفعل "أخبر" الإطارَ (NP NP S)، الذي يعني (فاعل Subject، ومفعول Object، وجُمَيلة Clause)، وأن الفعل "وجد" ليس له هذا الإطار، ولكن له الإطار (NP NP) –مثلًا– أمكننا أن نُلحق عبارة "المتفق عليه" بالفعل "أخبر" في الجملة الأولى، وبـ "المكان" في الجملة الثانية، وذلك على النحو التالي:

أخبرت [الزوج] [أن الطفل في المكان المتفق عليه].

وجدت [الزوجة] [المكان [المتفق عليه]].

ومن أشيع إطارات التصنيف الفرعي:

- الفعل اللازم. (عبارة اسمية NP) [فاعل]. انصرف الرجل.
- الفعل المتعدي. (عبارة اسمية NP) [فاعل]، (عبارة اسمية NP) [مفعول به]. كتب الطالب الدرس.
- الفعل المتعدي المتبوع بعبارة جر. (عبارة اسمية NP) [فاعل]، (عبارة اسمية NP) [مفعول به]، (عبارة جر PP). وضعتُ الكتاب على المنضدة.
- الفعل المتعدي إلى مفعولين. (عبارة اسمية NP) [فاعل]، (عبارة اسمية NP) [مفعول به مباشر] (عبارة اسمية NP) [مفعول به غير مباشر]. أعطى محمدٌ عليًّا كتابًا.
- الفعل اللازم المتبوع بعبارة جر. (عبارة اسمية NP) [فاعل]، (عبارة جر PP). اعتلى الإمام مؤيديه.
- الفعل المتبوع بجُمَيلة. (عبارة اسمية NP) [فاعل]، (جُمَيلة Clause). أعرف أنك مجتهد.
- الفعل المتعدي المتبوع بجُمَيلة. (عبارة اسمية NP) [فاعل]، (عبارة اسمية NP) [مفعول به]، (جُمَيلة Clause). أخبرته أنه ناجح.
-

ويتشابه هذا المصطلح مع مصطلح التكافؤ Valency الذي قدَّمه اللغوي الفرنسي لوسيان تيسنير Lucien Tesniere، والذي كان له تأثير بالغ في تطوير نماذج نحو الاعتماد Dependancy grammar في أوروبا وأمريكا. وهو مصطلح مشتق من الكيمياء، ويُستخدم في اللغويات للإشارة إلى عدد الروابط التي تكوِّنها العناصر النحوية مع بعضها، وأنواعها، فكما يمكن لعنصر ما أن يُكوِّن تكافؤات مختلفة في سياقات مختلفة، يعرض نحو التكافؤ Valency grammar نموذجًا للجملة يحتوي على عنصر أساسي –عادة الفعل- وعدد من العناصر الاعتمادية (ويشار إليها بطرق عدة منها المحمولات، والتعبيرات، والمكملات، والمتكافئات) يتحدد عددها وأنواعها بالتكافؤ المرتبط بالفعل[6]. وعلى هذا، يمكن أن يشتمل تكافؤ الفعل على عنصر واحد هو الفاعل، ومن ثم يعد هذا الفعل أحادي التكافؤ Monovalent، وقد يشتمل على عنصري الفاعل والمفعول به المباشر، ومن ثم هو ثنائي التكافؤ Bivalent. وهناك أفعال ثلاثية التكافؤ كالفعل "أعطى"، والفعل "وَضَعَ"، بيد أن التكافؤات التي يتحكم فيها الفعل أعطى (فاعل – مفعول به مباشر – مفعول به غير مباشر) تختلف عن التكافؤات التي يتحكم فيها وَضَعَ (فاعل – مفعول به مباشر – ظرفي مكاني)[6].

كما تلتقي فكرة الإطار التصنيفي مع تقسيم الفعل عند النحاة العرب من جهة المفعول؛ فهناك الأفعال اللازمة، وهي الأفعال التي تكتفي بفاعلها، ولا تقع على مفعول به؛ وهناك أفعال 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. وبتطبيق فكرة الإطارات التصنيفية على هذا التقسيم، يمكن ملاحظة بعض الإطارات التصنيفية:

الجدول رقم (2)
بعض إطارات التصنيف الفرعي في ضوء تقسيم الفعل بحسب المفعول به

نوع الفعل	الإطار	الوظيفة	الفعل	مثال
لازم	NP	فاعل	عاد	عاد المسافر.
لازم	NP PP	فاعل – فضلة	مات	مات المريض في المستشفى.
متعدٍّ لمفعول واحد	NP NP	فاعل – مفعول	قاتل	قاتل الجندي العدو.
متعدٍّ لمفعولين أصلهما المبتدأ والخبر	NP NP NP	فاعل – مفعول – مفعول	علم	علمت الطالب مجتهدًا.
متعدٍّ لمفعولين ليس أصلهما المبتدأ والخبر	NP NP S	فاعل – مفعول – مفعول	أخبر	أخبر المعلم التلميذ أن إجابته سديدة.
متعدٍّ لثلاثة مفاعيل	NP NP NP NP	فاعل – مفعول – مفعول – مفعول	أنبأ	أنبأتك السماء صافية.

ويكون إعراب الأمثلة –وفقًا لهذه الإطارات التصنيفية- على النحو التالي:
عاد [المسافر].
مات [المريضُ] [في المستشفى].
قاتل [الجنديُ] [العدوَ].
يعلم [اللهُ] [أنني صادق].
علم-[ت] [الطالبَ] [مجتهدًا].
أخبر [المعلم] [التلميذ] [أن إجابته سديدة].
أنبأ[تـــ]–[ك] [السماء] [صافية].

2) الوظائف الدلالية

يشير مصطلح الوظيفة الدلالية إلى العلاقة الأساسية بين المسند والكيانات التي ترد معه وتتعلق به، وتُكسبه معناه، ويطلق عليها محمولات؛ إذ يرتبط الفاعل (العبارة الاسمية)، ومكونات العبارة الفعلية بالفعل بطرق متنوعة؛ فالعبارة الاسمية (الطفل) في جملة "شرب الطفل اللبن" تسمى "الفاعل Agent" لفعلِ "الشرب"، في حين تسمى العبارة الاسيمة "اللبن" "الموضوع Theme"، وهو ما يقع عليه الفعل؛ أي أن جزءًا من معنى الفعل أن فاعله هو الفاعل.(*)

شرب [Predicate] الطفل [Agent] اللبن [Theme].

أما العبارات الاسمية داخل الجمل التي يكون رأسها الفعل "وَضَعَ"، فتُمَثَّل ثلاث علاقات: الأولى هي الفاعل، والثانية هي الموضوع، والثالثة الهدف goal؛ ففي جملة "وَضَعَ الطفل الكتاب على الطاولة" تكون العبارة الاسمية "الطفل" هي الموضوع، كما في المثال السابق، في حين تمثل العبارة الاسمية "الكتاب" موضوعًا، أما عبارة "الطاولة" فهي الهدف، وتُفْهَم الجملة كاملةً على أن فاعل الفعل "وَضَعَ" غيَّر مكان الموضوع إلى الهدف.

وضع [Predicate] الطفل [Agent] الكتاب [Theme] على الطاولة [Goal].

ويمكن التعبير عن معرفة المتكلم بالأفعال في ضوء علاقتها بعناصر الجملة بطريقة شكلية تُعَيِّن العلاقات الدلالية بين الفعل والفاعل والمفاعيل المنتقاة:

شرب(V) → NP, NP (Agent, Theme)
وضع(V) → NP, NP, PP (Agent, Theme, Goal)

تمثل هذه العلاقات الدلالية بعضًا مما يطلق عليه "الوظائف الدلالية Semantic roles"، وليس ثمة اتفاق عام على عدد هذه الوظائف. يتضمن الجدول التالي أهم هذه الوظائف الدلالية وأشيعها، مع توصيفها، والتمثيل لها.

يشير جورافيكسي ومارتن إلى وجود صعوبة تتعلق بالوصول إلى مجموعة نموذجية لهذه الوظائف، وصعوبة مماثلة تتعلق بإنتاج تعريف منهجي لأدوار، مثل: الفاعل، والموضوع، والأداة[10]. ولا يمنع ذلك من الإفادة من الوظائف الدلالية المتفق عليها -على مستوى التعريف أو العدد- في تمثيل المعنى.

(*) يُقصد بالفاعل الأول المصطلح النحوي Subject، أما الفاعل الثاني فهو مصطلح دلالي من المصطلحات المستخدَمة في النظرية الدلالية؛ ترجمة للمصطلح الإنجليزي Agent.

الجدول رقم (3)
أهم الوظائف الدلالية وأشيعها

الوظيفة الدلالية	التوصيف	مثال	ملاحظات
الفاعل AGENT	القائم بالفعل إراديًا.	شرح **المعلم** الدرس.	
الموضوع THEME	الشيء المتأثر بالفعل.	كسر الولد **الزجاج**.	يطلق عليه أيضًا PATIENT
الأداة INSTRUMENT	الوسيلة التي ينفذ بها الفعل أو الحدث.	أصاب الضابط اللص **بسلاحه الناري**.	
المُدْرِك EXPERIENCER	الشيء الذي يكون على وعي بالحدث الذي يصفه المسند، لكنه لا يكون متحكمًا فيه.	أصاب **الرجل** دوار.	
القوة FORCE	المثير غير الإرادي للفعل.	اقتلعت **العاصفة** الأشجار.	يطلق عليها أيضًا CAUSATIVE
المستفيد BENEFICIARY	الشيء المستفيد من الفعل المُنفَّذ.	من مكارم الأخلاق أن يساعد الغني **الفقير**.	
المكان LOCATION	المكان الذي يحدث فيه شيء أو فعل ما.	اندلعت مظاهرات حاشدة **بميدان التحرير**.	
الهدف GOAL	المكان أو الشيء الذي يتحرك نحوه شيء ما.	وصل الرئيس إلى **دبي**.	
المصدر SOURCE	المكان أو الشيء الذي يتحرك شيء ما منه.	أقلعت الطائرة من **القاهرة**.	
النتيجة RESULT	نهاية الحدث.	شيدت دبي العقارية **برجًا ضخمًا**.	

3) القيود (الاختيارات) الانتقائية

تُفَضِّل الأفعال –في الأغلب- محمولات من نوع محدد، وتُعْرَف هذه التفضيلات بالقيود الانتقائية، أو الاختيارات الانتقائية Selectional preferences، فالفعل "يأكل" -مثلًا- يميل إلى مفاعيل تعبر عن أصناف الطعام، وفواعل الفعل "يفكر" تتجه إلى البشر، في حين يميل الفعل "يعوي" إلى فاعل معين (الكلب).

على الرغم من شيوع مصطلح القيود الانتقائية في الدراسات الدلالية، فإن ثمة اتجاهًا نحو استخدام مصطلح الاختيارات الانتقائية؛ إذ يبدو الأخير –على حداثته- أكثر انضباطًا في دلالته على الظاهرة؛ لأن الاختيارات يمكن تجاهلها في الاستعارات، والمجاز[12]. ففي جملة نحو "صَرَّحَت القاهرة..." يقتضي الفعل "صَرَّحَ" فاعلًا بشريًا، وهو ما تجاوزته الجملة، وقد يأخذ الفعل "يأكل" -مثلًا- محمولات لا علاقة لها بالطعام، كما في جملة "تأكل النار الحطب".

وإذا كانت الوظائف الدلالية إحدى الطرق التي يُعَبَّر بها عن دلالات المحمولات في ضوء علاقاتها بالمسند، فإن القيود الانتقائية تُمَثِّل طريقة أخرى للتعبير عن المحددات التي يفرضها فعلٌ معين على أنواع المفاهيم المسموح لها بشغل دور المحمول[10].

وترتبط أهمية القيود الانتقائية في معالجة اللغات الطبيعية ارتباطًا أساسيًا بمشكلة الالتباس المعجمي والنحوي؛ إذ لم يحاول علم اللغة النظري أن يحدد طريقة يمكن من خلالها تطبيق نموذج كفاءة Competence model للتعامل مع هذه المشكلة عبر عمليات مقنعة –على المستوى النفسي- سواء في إدراك اللغة أو إنتاجها[13]. كما لم يثبت بعدُ ما إذا كان المعرب البشري يستخدم أسلوب التعقب الخلفي أم التوازي في التعامل مع الالتباس، أو ربما يستخدم استراتيجية الخلط، أو استراتيجية أخرى بديلة. وما زال هذا الأمر يمثل تحديًا لعلماء اللغة النفسية، ويبدو اكتشافه من الصعوبة بمكان؛ لأن الناس يتعاملون –عادة- مع الالتباس، ويصلون بسهولة إلى المعنى المقصود، وليست هناك أدلة كثيرة على أنهم يبذلون مجهودًا كبيرًا إضافيًا للتعامل مع المعاني الإضافية الممكنة[8].

ومن ثم كان على اللغويين الحاسوبيين إيجاد وسيلة فعالة لاختيار التحليل المناسب من بين تحليلات عدة ممكنة، وعلى هذا ارتبطت أهمية القيود الانتقائية بترتيب الإعرابات الممكنة للجملة؛ فتأخذ الإعرابات التي يأخذ فيها الفعل محمولات "طبيعية" ترتيبًا أعلى من المحمولات "غير القياسية المطردة". وهذا الأسلوب يسمح بالاختيار الجيد من بين الإعرابات. ويمكننا أن نسوق هنا المثال الذي عرضه جورافيكسي ومارتن[10]:

I want to eat some place that’s close to ICSI.

الترجمة: "أريد أن آكل مكانًا ما يكون قريبًا من آي سي إس آي."

ثمة احتمالان لإعراب هذا المثال وتفسيره دلاليًا؛ ففي حالة التفسير المقبول للفعل "آكل" باعتباره لازمًا، تكون عبارة "مكانًا ما يكون قريبًا من آي سي إس آي" مُتمِّمةً للإشارة إلى مكان حدث الأكل. أما في حالة التفسير غير المقبول باعتبار الفعل متعديًا، فتكون هذه العبارة مفعولًا به مباشرًا، تمامًا كالعبارة الاسمية "طعامًا ماليزيًا" في جملة:

I want to eat Malaysian food.

أريد أن آكل طعامًا ماليزيًا.

كيف يتسنى لنا أن نعرف أن عبارة "مكانًا ما يكون قريبًا من آي سي إس آي" ليست مفعولًا به مباشرًا في هذه الجملة؟ إن أحد المفاتيح المفيدة لمعرفة ذلك يكمن في معرفة الاختيارات الانتقائية للفعل "يأكل" الذي تجنح محمولاته إلى الأشياء المأكولة. وتقترح المجموعة الاستشارية ترميز الإطار الدلالي كما في المثالين الموضحين بالشكل رقم (9).

Entry: سَارَ
[ARG: arg1: [SELECTIONAL-RESTR: animate
THEMATIC-ROLE: agent]
Arg2: [SELECTIONAL-RESTR: location
THEMATIC-ROLE: destination]
PRED-CLASS: movement
SEM-REPT: 'arg1 CAUSE arg1 GO TO arg2 MANNAR walking']

Entry: أعطى
[ARG: arg1: [SELECTIONAL-RESTR: human
THEMATIC-ROLE: agent, source]
Arg2: [SELECTIONAL-RESTR: not (animate)
THEMATIC-ROLE: theme]
Arg3: [SELECTIONAL-RESTR: animate
THEMATIC-ROLE: destination]
PRED-CLASS: transfer of possession
SEM-REPT: 'arg1 CAUSE arg2 GO TO arg3']

الشكل رقم (9)
ترميز الإطار الدلالي للفعلين: سار، وأعطى.

5 الخاتمة

تمثل المعلومات الدلالية عامًلا مساعدًا لكثير من نظم المعالجة الآلية للغات الطبيعية، وقد تناولت هي الدراسة مجموعة الظواهر والمعلومات الدلالية التي ترمز في المعاجم والمدونات النصية المصممة لأغراض تكنولوجيا اللغة. واعتمدت الدراسة معايير المجموعة الخبيرة في هندسة اللغة مرجعًا لها، على مستوى المعلومات وكذلك طريقة ترميزها، تطبيقًا على بعض مداخل المعجم الحاسوبي العربي.

قائمة المراجع

[1] فايد، وفاء كامل. (2007). المتطلبات اللغوية لمعالجة التعابير الاصطلاحية العربية معالجة حاسوبية. *الندوة الدولية الأولى للحاسب واللغة العربية ISCAL، الرياض.*

[2] Attia, M. et al. (2011). Lexical Profiling of Arabic. *Proceedings of eLex*, pp. 23-33.

[3] Brisco, T. & Carroll, J. (1997). Automatic Extraction of Subcategorization from Corpora. *In the Proceeding of the 5th ACL Conference on Applied Natural Language Processing* . pp. 356 – 363.

[4] Carroll, J. (1998). Automatic Acquisition of Subcategorization Frames and Selectional Preferences from Corpora. *Talk given at the workshop of "Practical Acquisition of Large Scale Lexical Information" at CSLI, Stanford*, on April 23, 1998.

[5] Cruse, D. (1986). *Lexical Semantics*. Cambridge University Press.

[6] Crystal, David. (2008). *A Dictionary of Linguistics and Phonetics.* Sixth Edition. Blackwell Publishing Ltd.

[7] EAGELS. (1999). *Preliminary Recommendations on Lexical Semantic Encoding.* Final Report.

[8] Fromkin, V., Rodman, R., & Hyams, N. (2003). *An Introduction to language*. Wadsworth.

[9] Habash, N. (2010). *Introduction to Arabic Language Processing*. Morgan & Claypool Publishers Series.

[10] Jurafsky, D. & Martin, J. (2008). *Speech and Language Processing: an Introduction to Speech Recognition, Computational Linguistics and Natural Language Processing*, 2nd ed., Prentice Hall.

[11] Kearns, K. Lexical Semantics in Aarts, B. & Mcmahon, A. (2006). *The Handbook of English Linguistics*. Blackwell Publishing Ltd.

[12] Manning, C. & Schutze, H. (2000). *Foundations of Statistical Natural Language Processing*. The MIT Press Cambridge; Massachusetts; London; England.

[13] Wagner, A. & Mastropietro, M. (1996). Collecting and Employing Selectional Restrictions. *In Papers of the First Swiss-Estonian Student Workshop on Computational and Theoretical Linguistic*, Tartu, September 1996.

السيرة الذاتية:
أ.د/ وفاء كامل فايد

- أستاذة بكلية الآداب جامعة القاهرة.
- عضو مراسل بمجمع اللغة العربية بدمشق.
- خبيرة بمجمع اللغة العربية بالقاهرة.
- نشرت 53 بحثا باللغتين العربية والانجليزية، وأشرفت على 48 رسالة للماجستير والدكتوراه.
- حصلت على جائزة جامعة القاهرة التشجيعية عام 2004، وجائزة جامعة القاهرة التقديرية عام 2013.

- ألّفت وترجمت 9 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).

يوسف محمد يوسف أبو عامر

مدرس مساعد بقسم اللغة العربية بكلية الآداب، جامعة القاهرة. حصل على الماجستير في اللغويات الحاسوبية بتقدير ممتاز، مع التوصية بالطبع والتبادل، كما حصل على ليسانس الآداب من قسمي اللغة العربية وآدابها واللغة الإنجليزية وآدابها بكلية الآداب، جامعة القاهرة، وكذلك على دبلوم الترجمة الإنجليزية من قسم اللغة الإنجليزية. صمم محللًا صرفيًّا للغة العربية باستخدام محولات الحالة المحدودة Finite State Transducers، وله منشورات عن المعجم الحاسوبي العربي المصمم لتطبيقات المعالجة الآلية للغة العربية. شارك –مع آخرين– في ترجمة عدد من المقالات التي تهدف إلى إعطاء القارئ العربي مقدمة عن تحليل الخطاب. كما وضع نواة لمعجم ثنائي اللغة إنجليزي – عربي لترجمة المصطلحات المستخدمة في حقل اللسانيات الحاسوبية.

Symbolisation of Semantic Phenomena in Computational Arabic Lexicon

Wafaa Kamel Fayed[1] , Yousef Abou Amer[2]
Faculty of Literature - Cairo University
[1]wafkamel@yahoo.com
[2]ysfaboamer@gmail.com

Abstract: ***Semantics (the study of meaning) is a real challenge for researchers in Arabic computational linguistics as it involves a number of features that create a great deal of complexity such as synonymy, antonym, hyponymy, etc. Moreover, the richness of Arabic morphology and the absence of diacritization result in a high frequency of homographs and consequently lexical ambiguity.***
However, several text-processing applications such as information retrieval, text summarization and question answering systems require information related to the meaning of words. Therefore, lexical resources have to provide systems with semantic such as synonymy, meronymy and hierarchical relations in order to analyze sentences properly.
In this paper we aim at applying the standardizations of EAGlES on lexical semantic encoding for natural language processing applications to entries of Arabic computational lexicons. According to EAGLES, Semantic features are partitioned into three priority bands: high, medium and low according to their importance to applications.
In this paper we focus on high priority features as they are used in most of the applications and enabling technologies reviewed and are likely to be utilized in language technology products.