

المعالجة الآلية للتعديد الدلالي: أفعال الحركة نموذجاً

تستند التطبيقات في المعالجة الآلية للغة، مثل استخراج المعلومة وإعادة بنية وثيقة (document restructuring)، إلى قواعد معارف معجمية لتأويل نص معين. فهذه القواعد ترمز معاني الكلمات وما يرتبط بها من تركيب وصرف. فمثلاً قاعدة WordNet تعد القاعدة الأكثر شيوعاً بالنسبة للغة الإنجليزية (انظر تطبيقات فيلبوم (1998) Felbaum ولغات أخرى. غير أن هذه القاعدة تعاني من مشكل أساسي يتمثل في كونها تقوم على تعداد المعنى / المعنى التصريحي (sense enumeration). وهذه الخاصية تطرح مشاكل متعددة بالنسبة لتطبيقات المعالجة الآلية للغة على مستوى التأويل المنتظر. وهذا يعود، أساساً، إلى عيوب WordNet التالية:

1. عدم كفاية أو الإفراط في تبسيط التمثيل المعرفي
2. تغطية غير كافية في تطبيق المجال (بخصوص توزيع وتحديد المجالات انظر WordNet)
3. ضعف دعم تأويل التصورات الجديدة، يعني التصورات غير المدونة في القاعدة ومن ثم الاستعمالات الجديدة غير المدونة بالنسبة للتصورات المدونة.

تجنباً لهذه المشاكل والثغرات نتبنى في معالجتنا للتعديد الدلالي آلياً مقاربة حديثة لقواعد المعارف المعجمية ألا وهي المعجم المولد (بوستوفيسكي (1995-1998-2001) *Generative Lexicon* (Pustejovsky). نقترح في هذه الورقة تمثيلاً معرفياً للأنطولوجيات اللغوية الفرعية والتي هي في الواقع أنطولوجيات لغوية مستلهمة من الدلالة المعجمية لتأويل النص في مجال مخصوص. سنركز في هذه الورقة على معالجة ظاهرة التعدد الدلالي ونأخذ أفعال الحركة نموذجاً لها. فالملاحظ أن هناك استعمالات مختلفة لأفعال مثل دفع ودخل ومر وعبر كما في البنائات التالية:

4. (أ) دخل إلى البيت
- (ب) دخل في أزمة
- (ج) دخل في غيبوبة
5. (أ) دفع الرجل العربية
- (ب) دفعه القرار الجائر إلى الانتحار
6. (أ) مر على باب الحديقة
- (ب) مر بظرف صعبة

سنبين كيف أن العلاقات الدلالية للأنواع (qualia)، وليس المعنى التصريحي، تمكن من ضبط العلاقات المتعارضة الموجودة بين الوحدات الدلالية كيفما كان موقعها في الأنطولوجيا. وهكذا، فإن ضبط تعدد المعنى يقوم على البحث على المعلومة بواسطة استغلال رفع التباس معاني هذه الأفعال على أساس التصنيف الأنطولوجي والسمات الدلالية. وهذا يمكن من الفهم الآلي للنصوص الذي يستند إلى معطيات ترتبط بالمساهمة الدلالية للكلمة في الجملة أو السياق. وهكذا يتم رفع الالتباس المعجمي بفضل الغنى الواسع للمعلومة المضمنة في وصف الوحدات المعجمية والقيود على أنماط الموضوعات التي تقبلها في السياق.