**Semantic relations in compound nouns: Perspectives from inter-annotator agreement**

Prabha Yadava, Elisabetta Jezekb, Pierrette Bouillonc, Tiffany J. Callahana,

Michael Badaa, Lawrence E. Huntera, and K. Bretonnel Cohena

**a** Computational Bioscience Program, University of Colorado School of Medicine, Aurora, Colorado 80045 USA

**b** Department of Humanities, University of Pavia, Italy

cFaculté de Traduction et d’Interprétation, Université de Genève, Switzerland

Abstract

Semantic relations have been studied for decades without yet reaching consensus on the set of these relations. However, biomedical language processing and ontologies rely on these relations, so it is important to be able to evaluate their suitability. In this paper we examine the role of inter-annotator agreement in choosing between competing proposals regarding the set of such relations. The experiments consisted of labeling the semantic relations between two elements of noun-noun compounds (e.g. cell migration). Two judges annotated a dataset of terms from the biomedical domain using two competing sets of relations and analyzed the inter-annotator agreement. With no training and little documentation, agreement on this task was fairly high and disagreements were consistent. The results support the utility of the relation-based approach to semantic representation.

Keywords:

MeSH Unique ID: D012660

Tree Number(s): L01.143.506.598.745

MeSH Heading Natural Language Processing

Tree Number(s) L01.224.050.375.580

MeSH Heading Evaluation Studies as Topic

Tree Number(s) E05.337, N05.715.360.335.

Introduction

Linguists have tried to discover the basic building blocks of semantic relations between nouns for decades, but there is still little consensus about what the set of those building blocks might be. This is an important problem for biomedical natural languge processing and biomedical ontologies, because those building blocks are at the heart of our information extraction tasks and the structure of our ontologies. Compound nouns are crucial to practical tasks like knowledge representation and to theoretical problems like understanding compositionality in semantics [7; 20]. Compound nouns are formed by a sequence of two or more nouns [3]. In writing, they may appear as two tokens (*knockout mouse*), a hyphenated word (*nucleotide-excision*), or a single token (database) [3]. They are about twice as common in written English as they are in spoken English (248/million words in newswire text versus 123/million words in conversation)[3]. Linh (2010) [12] reviews a number of studies of the incidence of compound nouns in technical texts, reporting that one study found that 27% of words in scientific abstracts were in compound nouns, another study found that 11.86% of anaphors are compound nouns, and another found that 15.37% of a technical corpus was made up of compound nouns.

The study of compound nouns dates back to Pānini and Kātyāyana and Patañjali [9], but an enormous amount of work remains to be done, particularly the semantics of the relations between the nouns in a compound, and they remain the topic of a considerable amount of research in both linguistics and natural language processing [13]. Various authors have attempted to describe the relation between the elements of compound nouns from a theoretical perspective [18]. Likewise, a number of studies have attempted to analyze compound nouns automatically [6; 10; 19; 21; 22]. All of these studies are based on specific representations of the relations that can hold between the nouns in a compound. This raises a question: can humans reliably label the relation that holds between any two specific nouns in any specific compound? If they cannot, then we must question the validity of the representation itself, and we must consider the possibility that any principled investigation of the relations in compounds, whether from a theoretical or a practical perspective, is impossible. On the other hand, if they can, then it might be possible to train computers to do the same task, which could enable considerable advances in natural language processing.

We can address the reliability of labeling through examining inter-annotator agreement when two or more analysts label the relations in a sample of compound nouns. However, we are not aware of any studies that have looked at inter-annotator agreement in compound nouns. Identifying relations in compound nouns, whether done by humans or by computers, is a non-trivial task because there is an enormous amount of ambiguity in the correspondence between semantics and syntactic structure. For example, Table 1 gives a number of examples using the biomedical term *forceps*. We note that *forceps* can exist in at least five relations with another noun in a compound—that is to say, the same noun-noun syntactic structure can correspond to at least five relations between the first noun and *forceps*.

*Table 1: Identical syntactic structures can reflect a wide variety of semantic relations*

|  |  |
| --- | --- |
| **Relation** | **Example** |
| Used\_on | *Bone forceps* |
| Instrument\_for | *Epilation forceps* |
| Shape\_of | *Mosquito forceps* |
| Operated\_by | *Thumb forceps* |
| Named\_for | *Kelly forceps* |

This article investigates the ability of humans to reliably label semantic relations between the elements of noun compounds. We hypothesize that inter-anotator agreement on labelling the semantic relations in compound nouns is a useful indicator of the validity of proposed set of semantic relations and can be used in choosing between two competing theories. A first step is to understand the relation between the nouns that make up compound nouns (e.g., protein kinase or cell migration). Understanding this relation is one of the most difficult problems in language processing [11; 13; 16], however, it is crucial to do because compound nouns represent an enormous amount of knowledge about our conceptions of the world and the kinds of things that interact within it. The hypothesis was tested using two sets of relations based on two different linguistic theories and with two different contexts of theoretical status and computational applications.

Materials

Generative Lexicon

The first set of relations is based on the Generative Lexicon (GL) theory [15]. The theory is an attempt to explain how compositionality contributes to lexical semantics. Bouillon et al., 2012 [4] posit two basic elements of lexical semantic representations: Qualia and/or Argumental. This distinction is somewhat problematic, given the fact that qualia relations involve predicates and their arguments as well. We will re-visit this issue in the discussion. They identified four basic Qualia relations: Formal, Constitutive, Telic, and Agentive. We used the set of Generative Lexicon relations described in Bouillon et al., 2012 [4]. These relations are meant to be general and elementary, embodying a hypothesis about the fundamental building blocks of semantic representations.

Rosario and Hearst

Rosario and Hearst (2001) [16] identified 38 relations roughly following the linguistic theories of Levi and Warren, but the motivation for the relations is less language-theoretic and more application-oriented. Specifically, it is oriented towards representing the semantics and knowledge structures of biomedical literature. The expectation is that these relations would work better than the Bouillon et al., 2012 [4] tags as our study is based on the biomedical domain. The set of relations proposed by Rosario and Hearst is meant to be domain-specific and non-elementary. In comparison to many other proposed sets of semantic relations, including that of Generative Lexicon theory (Table 2), the Rosario and Hearst relations do not posit a set of semantic primitives embodying a knowledge representation schema that is specifically tailored to biomedical science without making any claims about what relations might exist in other domains.

*Table 2: Summarizes the characteristics of the two sets of relations in terms of their size, goals, and generality/domain specificity*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Relation set** | **Size** | **Goal** | **General** | **Domain-specific** |
| Generative Lexicon | 5 | Theoretical | x |  |
| Rosario and Hearst | 38 | Application-oriented |  | x |

The sample of terms for annotation was drawn from the Gene Ontology (GO). The GO is a good test case for this study because its content is clearly relevant to the biomedical domain and because it contains a large number of noun compounds. The GO project (http://geneontology.org/) was founded in 1998. Open access, the GO is a community based effort to create unified representation of gene and gene products in a species independent manner. The GO invites contributions from the research community and new terms and annotations are added every day. It provides an ontology of defined terms representing gene product properties. The ontology covers three domains namely Cellular Component, Molecular Function, and Biological Process. Each GO term includes a term name, term accession number, a namespace indicating the domain to which it belongs, and a cited textual description [1]. As of September 2014, 41,775 GO terms were reported (Gene Ontology Consortium, 2015) [5].

To select the sample, all of the terms in the GO were tagged with their part of speech using the CLEAR suite of language processing tools [2; 14] (Tagging errors are noted in the data set). All terms with exactly two words, such that both words are nouns—that is, compound nouns—were pulled from the full set. We then selected a random sample of 101 words from the compound nouns (intending 100, with an extra in case of tagging errors).

To specifically understand the meaning of a GO term, we have to understand the entity relations. For example, to understand the meaning of the term *chondrocyte development*, we have to understand that the relation between the entites is an activity/physical process (development) that is undergone by the chondrocytes. The biological domain includes a great number of technical terms and it becomes difficult to understand them without knowing the exact meaning of the entities and the relation between them. For example, *motor activity* and *thrombin activity* have identical syntactic structures, but the relationships between them is different: resulting in motion in the first and activity of thrombin in the second. Similarly, *chaperone activity* and *lysozyme activity* both appear to be similar, but *chaperone activity* assists in the correct non-covalent assembly of polypeptide-containing structures in vivo, but is not a component of these assembled structures when they are performing their normal biological function. Whereas, *lysozyme activity* catalysis the hydrolysis of the beta-(1->4) linkages between N-acetylmuramic acid and N-acetyl-D-glucosamine and is thus destructive to the cell walls of certain bacteria.

Methods

Annotators

One annotator has a bachelor’s degree in biology. The other annotator is a cardiovascular technologist with a doctorate in linguistics. They were instructed to base their annotations of the relation in a term on the definition of that term on the GO web site.

Evaluation

The main motivation behind this study was to examine whether GO terms can be represented effectively using a specific set of tags and their relations and if the annotators are able to agree on the kind of relations that exists between the compound nouns. We tried to answer the above question based on the Generative Lexicon relations (Table 3) and the Rosario and Hearst relations (Table 4). The inter-annotator agreement was measured using the F-measure and the Cohen’s Kappa value.

Results

*Table 3: Bouillon et al., Relations and corresponding ontology relations with examples*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Categorical Relations** | **GL Relations** | **Descriptions** | **Ontology Relations** | **Examples** |
| Formal | is\_a | N2 is a kind of N1 | is\_a/ subclass of | *predatory behavior* |
| shape\_of | N1 has the shape of N2 |  |  |
| holds | N1 holds N2 | contains |  |
| Constitutive | made\_of | N1 is made of N2 | has\_proper\_part | *dynactin motor* |
| part\_of | N1 is a part of N2 | proper\_part\_of |  |
| located\_in | N1 is spatially/temporally located in N2 | located\_in |  |
| member\_of | N1 is member of N2 | member\_of | *kinin cascade* |
| has members | N1 has N2 as members | has\_member |  |
| Telic | predicate | N1 has the purpose of (Predicate) N2 | has\_function |  |
| used\_for | N1 is used for the activity N2 |  | *chondrocyte differentiation* |
| aims\_at | N1 has N2 as result/end goal |  | *keratinocyte development* |
| played\_by | N1 denotes the function which is N2. | has\_function | *cholinesterase activity* |
| Agentive | caused\_by | N1 is created/brought into existence/caused by N2 |  | *heart wedging, chondrocyte hypertrophy* |
| derived\_from | N1 is derived/extracted from N2 | derives\_from, transformation\_of |  |
| Argumental | argument | N2 is an argument of N1 |  | *p53 binding* |

*Table 4: Rosario and Hearst semantic relations*

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Examples** | **Name** | **Examples** |
| Wrong parse | exhibit asthma, ten drugs | Time of (1-2) | morning headache, hour headache |
| Subtype | headaches migraine, fungus candida | Measure of | relief rate, asthma mortality, asthma morbidity |
| Activity/Physical process | bile delivery, virus reproduction, bile drainage, headache activity | Person/center who treats | headache specialist, headache center, diseases physicians, asthma nurse |
| Ending/reduction | migraine relief, headache resolution | Instrument (1-2) | aciclovir therapy, chloroquine treatment |
| Defect in Loc. | lung abscess, artery aneurysm | Instrument (2-1) | vaccine antigen, biopsy needle |
| Change | papilloma growth | Instrument (1) | heroin use, internet use, drug utilization |
| Produces (on a genetic level) | polyomavirus genome, actin mRNA, CMV DNA, protein gene | Object | bowel transplantation, kidney transplant, drug delivery |
| Cause (1-2) | asthma hospitalizations, AIDS death | Misuse | drug abuse, acetaminophen overdose |
| Cause (2-1) | flu virus, diarrhea virus | Subject | headache presentation, glucose metabolism |
| Characteristic | receptor hypersensitivity, cell immunity | Purpose | headache drugs, HIV medications |
| Physical property | blood pressure, artery diameter | Topic | time visualization, headache questionnaire |
| Defect | hormone deficiency, CSF fistulas | Location | brain artery, tract calculi, liver cell |
| Physical Make Up | blood plasma, bile vomit | Modal | emergency surgery, trauma method |
| Person afflicted | AIDS patient, BMT children | Material | formaldehyde vapor, aloe gel, gelatin powder |
| Demographic attributes | childhood migraine, infant colic, women migraineur | Frequency/time of (2-1) | headache interval, attack frequency, football season |
| Bind | receptor ligand, carbohydrate ligand | Activator (1-2) | acetylcholine receptor, pain signals |
| Research on | asthma researchers, headache study | Activator (2-1) | headache trigger, headache precipitant |
| Attribute of clinical study | headache parameter, attack study, headache interview | Inhibitor | adrenoreceptor blockers, influenza prevention |
| Procedure | tumor marker, genotype diagnosis | Standard | headache criteria, society standard |
| Beginning of activity | headache induction, headache onset |  |  |

**The Generative Lexicon tag set from Bouillon et al.**

The annotators used eight relations to annotate the 101 GO terms (see Table 5). The most commonly used relation by annotator 1 was *argument* followed with *played\_by*. The most commonly used relation by annotator 2 was *played\_by* followed with *used\_for*. Annotator 1 thought there were no proper relations available for four terms: *larval development*, *predatory behavior*, *lymphocyte anergy,* and *lymphocyte homeostasis*. Table 6 shows the results in terms of true positives (1 for each match between the two annotators), false positives (1 for each mismatch between the two annotators), and false negatives (also 1 for each mismatch between the two annotators), and the corresponding measures of inter-annotator agreement. Cohen’s Kappa value was 0.47 and the inter-annotator agreement, calculated as F-measure, was 0.58 The Cohen’s Kappa value indicates a fair/good level of reliability according to the Green scale (1997). The annotators agreed that the Telic relation was the most frequent relation, followed by the Argumental relation. Annotator 1 thought 54.45% of the terms were Telic and annotator 2 thought 70.29% of the terms were Telic (Table 5). 36.63% of the terms were annotated as Argumental by annotator 1 and 19.80% of the terms were annotated as Argumental by annotator 2.

*Table 5: Data based on Bouillon’s relations*

|  |  |  |  |
| --- | --- | --- | --- |
| **Relation** | **Annotator 1 (%)** | **Annotator 2 (%)** | **Examples** |
| Formal [is\_a] | 0 | 0.99 | *predatory behavior* |
| Constititive [made\_of] | 0.99 | 0.99 | *dynactin motor* |
| Constitutive [member\_of] | 0.99 | 0 | *kinin cascade* |
| Telic [used\_for] | 0 | 20.79 | *chondrocyte differentiation,* |
| Telic [aims\_at] | 21.78 | 18.81 | *translation reinitiation,* |
| Telic [played\_by] | 32.67 | 30.69 | *GTPase activity,* |
| Agentive [caused\_by] | 2.97 | 7.92 | *chondrocyte hypertrophy, heart wedging* |
| Argumental [Argument] | 36.63 | 19.80 | *protease binding, p53 binding* |
| Un-annotated | 3.96 | 0 | *lymphocyte homeostasis* |

*Table 6: Annotation data based on Bouillon’s tags*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **TP** | **FP** | **FN** | **P** | **R** | **F** | **κ** |
| 59 | 42 | 42 | 0.58 | 0.58 | 0.58 | 0.47 |

Table 8[[1]](#footnote-1) shows the confusion matrix for these relations. Note that disagreements between the two annotators are largely systematic. For example, Annotator 1 and Annotator 2 used the *aims\_at*relation a similar amount of times (22 and 19), agreeing in the case of 10 noun compounds (Annotator 1 10/22 (45.45%); Annotator 2 10/19 (52.63%)). Of the disagreements, 8/22 times that Annotator 1 labelled instances of the *aims\_at*relation, Annotator 2 labelled the *used\_for* relation; 7/19 times that Annotator 2 labelled the *aims\_at* relation, Annotator 1 annotated the *argument* relation. Thus, refining the guidelines such that it is clearer when to use *aims\_at* versus *argument* and *used\_for* would have a large effect on the inter-annotator agreement for all three of these relation types. This refinement would most likely bring the inter-annotator agreement, as calculated by Cohen’s Kappa, to 0.65 and as calculated by F-measure to 0.73 (Table 7). Again, this is due to the overall quite systematic nature of the disagreements between the two annotators.

The ontology relations corresponding to the Generative Lexicon relations are shown in Table 3. We observed that there were no corresponding ontology relations for most of the Generative Lexicon relations suggesting the need for additional work.

*Table 7: Recalculated F-measure and Cohen Kappa values*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **TP** | **FP** | **FN** | **P** | **R** | **F** | **κ** |
| 74 | 27 | 27 | 0.73 | 0.73 | 0.73 | 0.64 |

**Rosario and Hearst relations:**

The annotators used 10/38 of the Rosario and Hearst relations to annotate the 101 GO terms. The inter-annotator agreement, calculated via Cohen’s Kappa, was 0.38 (Table 9).

*Table 9: Annotation data based on the Rosario and Hearst relations*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **TP** | **FP** | **FN** | **P** | **R** | **F** | **κ** |
| 71 | 30 | 30 | 0.70 | 0.70 | 0.70 | 0.37 |

The inter-annotator agreement calculated using the F-measure was 70.29%. The maximum number of terms were annotated as *Activity*/*Physical Process* followed by *Characteristics* and *Material: Bind*. Annotator 1 annotated two terms as *Wrong Phrase* and felt there was no proper relation available for one GO term, namely *Kinin Cascade*. The annotators observed that there were no proper representation of movement terms (i.e., *Cilium Movement)*. A confusion matrix was created based on the annotated relations (Table 11). The annotators disagreed on terms like *Protease Binding*, *p53 binding* etc., which annotator 1 thought was *Material: Bind* and annotator 2 thought was *Characteristic*. The inter-annotator agreement is good given the fact that no training was provided. Table 9 shows the break-down in terms of true positives, false positives, and false negatives, as well as the corresponding values for Cohen’s Kappa and the F-measure.

Table 11 shows the confusion matrix for the results. Again, we see that the disagreements are quite systematic. Of the 30 disagreements, 20 (2/3) were from one cell of the table: Annotator 1 classified 20 compounds as having the *Material: Bind* relation, while Annotator 2 classified the same 20 compounds as having the *Characteristic* relation. Addressing this issue would bring the inter-annotator agreement to 0.77 if calculated as Cohen’s Kappa and to 0.90 if calculated as F-measure (Table 10).

*Table 10: Recalculated F-measure and Cohen Kappa values*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **TP** | **FP** | **FN** | **P** | **R** | **F** | **κ** |
| 91 | 10 | 10 | 0.90 | 0.90 | 0.90 | 0.77 |

Discussion

The Inter-annotator agreement was much higher for the Rosario and Hearst relations than for the Generative Lexicon relations, despite the fact that the set of Rosario and Hearst relations are much larger than the set of Generative Lexicon relations.

It's premature to say why this is the case, but we can propose some avenues for future investigation: (a) This result might be related to the fact that the Rosario and Hearst relations are domain-specific, while the Generative Lexicon relations are not; (b) This result might be related to the fact that the Generative Lexicon relations are abstract and theoretically motivated, while the Rosario and Hearst relations are concrete and motivated by practical considerations; (c) It might be related to the observation that the annotators only used 10 of the Rosario and Hearst and relations implying that the difference in size might not be as big as it seems and the difference in Inter-annotator agreement may not be quite as surprising; (d) The difference in Inter-annotator agreement might go away with actual annotation guidelines and training; and (e) We should also point out that the affordances of the two are different--in particular, the Rosario and Hearst relations might be better for defining information extraction tasks while the Generative Lexicon’s relations might be better for supporting inference.

Conclusion

The inter-annotator agreement in both cases approached that of many completed and published corpus annotation projects, even with very minimal guidelines and no real training. The agreement on this task was fairly high in both cases and disagreements were quite consistent, supporting the basic soundness of the relation-based approach to semantic representation and suggesting that it is not overly subjective. From a methodological perspective, the results suggest that higher levels of agreement and reliability can be reached with some training and refinement of the guidelines. Inter-annotator agreement was different between the two sets of relations, suggesting that inter-annotator agreement can differentiate between semantic representations, although a number of possible explanations for those differences should be pursued in future work.

The relatively high inter-annotator agreement suggests that the descriptions and examples of the relations in the Bouillon and Rosario and Hearst papers were easy to follow and that the annotators were able to clearly delineate the relations and the tags in most cases. This is consistent with the claim that they are precise and not overly subjective in their interpretability and applicability. Disagreements between the analysts were quite consistent. This suggests that a higher level of agreement and reliability can be achieved with a little training and refinment of the guidelines. This study should be replicated on a larger scale with proper guidelines and training to achieve a higher level of reliability.

An additional benefit of approaching the evaluation of a set of relations through an annotation task was that we uncovered some shortcomings of the relations. We noted that (a) there is no good representation of movement in the Rosario and Hearst relations and (b) some of the GO terms were not representable at all with those relations. In the case of the Generative Lexicon relations, we observed frequent confusion between qualia and argument (especially *used\_for* and *aims\_at*). This suggests that there is a need to clarify the demarcation between the two. A fruitful direction for future work would be to evaluate the nature of any correspondences that might exist between the two sets of relations. The work reported here contributes to the groundwork for such an effort.

Acknowledgements

We thank NIH grants LM008111 and LM009254 to Lawrence E. Hunter and NSF grant IIS-1207592 to Lawrence E. Hunter and Barbara Grimpe for funding this work.

References

[1] M. Ashburner, C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, and J.T. Eppig, Gene Ontology: tool for the unification of biology, *Nature genetics* **25** (2000), 25-29.

[2] S. Bethard, P.V. Ogren, and L. Becker, ClearTK 2.0: Design Patterns for Machine Learning in UIMA, in: *LREC*, 2014, pp. 3289-3293.

[3] D. Biber, S. Johansson, G. Leech, S. Conrad, E. Finegan, and R. Quirk, *Longman grammar of spoken and written English*, MIT Press, 1999.

[4] P. Bouillon, E. Jezek, C. Melloni, and A. Picton, Annotating qualia relations in Italian and French complex nominals, (2012).

[5] G.O. Consortium, Gene Ontology consortium: going forward, *Nucleic acids research* **43** (2015), D1049-D1056.

[6] C. Dima and E. Hinrichs, Automatic Noun Compound Interpretation using Deep Neural Networks and Word Embeddings, *IWCS 2015* (2015), 173.

[7] J. Fan, K. Barker, and B.W. Porter, The knowledge required to interpret noun compounds, in: *IJCAI*, Citeseer, 2003, pp. 1483-1485.

[8] R. Huddleston and G.K. Pullum, The Cambridge grammar of the English Language. *Cambridge: Cambridge University Press* (2002), 1-23.

[9] S.D. Joshi and J. Roodbergen, *Patañjali's Vyākaraṇa-Mahābhāṣya Sthānivadbhāvāhnika: P. 1.1. 56-1.1. 57*, Bhandarkar Oriental Research Institute, 1990.

[10] M. Lauer, Conceptual association for compound noun analysis, in: *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, Association for Computational Linguistics, 1994, pp. 337-339.

[11] M. Lauer and M. Dras, A probabilistic model of compound nouns, *arXiv preprint cmp-lg/9409003* (1994).

[12] N.M. Linh, Noun-noun combinations in technical English, (2010).

[13] P. Nakov, On the interpretation of noun compounds: Syntax, semantics, and entailment, *Natural Language Engineering* **19** (2013), 291-330.

[14] P.V. Ogren, P.G. Wetzler, and S. Bethard, ClearTK: A UIMA toolkit for statistical natural language processing, *Towards Enhanced Interoperability for Large HLT Systems: UIMA for NLP* **32** (2008).

[15] J. Pustejovsky, The generative lexicon, (1995).

[16] B. Rosario and M. Hearst, Classifying the semantic relations in noun compounds via a domain-specific lexical hierarchy, in: *Proceedings of the 2001 Conference on Empirical Methods in Natural Language Processing (EMNLP-01)*, 2001, pp. 82-90.

[17] B. Rosario, M.A. Hearst, and C. Fillmore, The descent of hierarchy, and selection in relational semantics, in: *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, Association for Computational Linguistics, 2002, pp. 247-254.

[18] D.O. Séaghdha, Learning compound noun semantics, *University of Cambridge, Cambridge, UK* (2008).

[19] S. Tratz and E. Hovy, A taxonomy, dataset, and classifier for automatic noun compound interpretation, in: *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, 2010, pp. 678-687.

[20] A. Tribble and S.E. Fahlman, Resolving Noun Compounds with Multi-Use Domain Knowledge, in: *FLAIRS Conference*, 2006, pp. 122-127.

[21] L. Vanderwende, Algorithm for automatic interpretation of noun sequences, in: *Proceedings of the 15th conference on Computational linguistics-Volume 2*, Association for Computational Linguistics, 1994, pp. 782-788.

[22] B. Verhoeven, W. Daelemans, M. Van Zaanen, and G. Van Huyssteen, Automatic Compound Processing: Compound Splitting and Semantic Analysis for Afrikaans and Dutch, *ComAComA 2014* (2014), 20.

**Address for correspondence**

Kevin Bretonnel Cohen, e-mail id : kevin.cohen@gmail.com

1. Table 8 and Table 11 are on the GitHub Repository: 'https://github.com/KevinBretonnelCohen/SemanticRelationsCompoundNouns.git/' [↑](#footnote-ref-1)