

Mapping WordNet Concepts with CPA Ontology

Svetla Koeva, Tsvetana Dimitrova, Valentina Stefanova, Dimitar Hristov

Institute for Bulgarian Language

Bulgarian Academy of Sciences

svetla, cvetana, valentina, dimitar@dcl.bas.bg

Abstract

The paper discusses the enrichment of WordNet data through merging of WordNet concepts and Corpus Pattern Analysis (CPA) semantic types. The 253 CPA semantic types are mapped to the respective WordNet concepts. As a result of mapping, the hyponyms of a synset to which a CPA semantic type is mapped inherit not only the respective WordNet semantic primitive but also the CPA semantic type.

1 Introduction

The paper presents/discusses an effort on enriching the data in WordNet and the links between WordNet concepts through expansion of the number of noun semantic classes through/by mapping the WordNet data (Miller et al., 1990) with the data in another resource – the Pattern Dictionary of English Verbs (PDEV) (Hanks, 2004; Hanks, 2005; Hanks, 2008).

WordNet synsets are classified into semantic primitives (also called semantic classes). Verbs and nouns are distributed into more elaborate classes (Miller et al., 1990), with corresponding labels (noun.person, noun.animal, noun.cognition; verb.cognition, verb.change, etc.) being assigned to them. The information about semantic primitives has been used in a number of efforts to verify/test and enrich semantic relations between noun and verb synsets (such as the type of morphosemantic relations – Agent, Undergoer, Instrument, Event, etc. – that link verb/noun pairs of synsets that contain derivationally related literals) (Fellbaum, 2009).

The semantic classification of WordNet nouns and verbs is consistent and useful for many language processing tasks. However, the natural language understanding and generation requires a precise and granular prediction for the set of concepts

that could saturate the arguments of a verb. Consider the verb {read:5} 'interpret something that is written or printed' and its sentence frame *Somebody —s something*. Obviously, not every noun classified as *noun.person* will/can be selected/collocate by/with the verb {read:5} as its subject and not every noun that is not classified as *noun.person* can be the object of the verb. Therefore, we assume that the WordNet noun semantic classes can be further specified in order to correlate more precisely with the verb-noun selecting requirements. To sum up, although the information is readily available in WordNet, not all useful information is explicitly accessible.

In this paper, we present an effort at mapping the WordNet concepts with the Corpus Pattern Analysis (CPA) semantic types that are part of the Pattern Dictionary of English Verbs (PDEV). PDEV is built on the basis of the lexicocentric Theory of Norms and Exploitations (Hanks, 2013) and exploits the CPA mechanism to map meaning onto words in text. PDEV consists of verb patterns and semantic types of their nominal arguments organized within the so-called CPA ontology.

Our goal is then twofold: to identify the concept or the set of concepts to which a given CPA semantic type corresponds and to explore the structures of the two hierarchies: WordNet semantic primitives and CPA semantic types.

The paper is organized as follows: in section 2, we present our motivation for the work before discussing different attempts at semantic classification of nouns in section 3. Section 4 briefly presents the CPA ontology, while section 5 outlines some issues with the WordNet noun hierarchy. The effort at mapping the CPA semantic types and WordNet concepts is discussed in section 6, with a comparison between the two structures in section 7 and some preliminary conclusions; our plans for future work are given in section 8.

2 Motivation

There are many examples, such as in (1) where the sentence frame in (1a) signals that the verb can have both human and non-human subject argument. Further, (1c), which has a definition comparable to (1a), leaves only non-human subject argument. In addition, the non-human subject arguments both in both (1b) and (1c) may both be specified as animate.

(1)

a. {purr:1, make vibrant sounds:1} 'indicate pleasure by purring; characteristic of cats'

Something —s; *Somebody* —s

b. {moo:1, low:4} 'make a low noise, characteristic of bovines'

Something —s

c. {meow:1, mew:1} 'cry like a cat; the cat meowed'

Something —s

Noun semantic primitives cannot be employed for detailed selectional restrictions on arguments because their organization is too general and some semantic classes can be missing or inappropriate. For example, the sentence frames in (2) do not specify that the verbs can be combined with nouns like *idea* (noun.cognition), *result* (noun.communication), *victory* (noun.event) but cannot co-occur with nouns such as *stone*, *table*, *sky*, etc.

(2)

{achieve:1, accomplish:2, attain:4, reach:9} 'to gain with effort'

Somebody —s *something*

Something —s *something*

Somebody —s *that CLAUSE*

To find a match between nouns and verbs, we hypothesize that verb hypernym/hyponym trees combine verbs with similar or equivalent semantic and syntactic properties.

Further, it can be tested whether verb synsets combine with noun classes that can be identified within the WordNet structure if a more detailed classification of nouns (which further specifies the semantic classes) – in line with the CPA semantic types ontology – is provided. Here, we present our work on mapping the WordNet concepts and the CPA semantic types.

Previous work on mixing resources and enriching the information on semantic and syntactic behavior of verbs encoded in WordNet builds upon resources – one or more than one – that use (Levin, 1993)'s verb classes (Dorr, 1997; Korhonen, 2002; Green et al., 2001). Proposals involve mixing up information from WordNet and Longman Dictionary of Contemporary English (Dorr, 1997; Korhonen, 2002); VerbNet (also based on Levin's classes) and FrameNet (Shi and Mihalcea, 2005); and VerbNet and PropBank (Pazienza et al., 2006). To the best of our knowledge, however, WordNet concepts and CPA ontology have not been mapped and compared yet, and below we propose such an effort.

3 Semantic classes of nouns

Although WordNet nouns are classified in a number of classes labeled by semantic primitives, numerous linguistic works argue that nouns have referential value and cannot be reduced to a set of primitives.

(Wierzbicka, 1986) claims that most (prototypical) nouns identify a certain kind of entity, a concept, but positively and not in terms of mutual differences. Thus, the function of a noun is to single out a certain kind of entity and its meaning cannot be reduced to any combination of features though it may be described using features.

In numerous works, (Wierzbicka, 1984; Wierzbicka, 1985) enumerates features such as shape, size, proportions, function, etc. that can be used in definitions of objects but in a semantic formula, these features have to be subordinated to a general taxonomic statement. For example, in conceptual representation of count/mass nouns, (Wierzbicka, 1988) motivates 14 classes of language terms, with each class being conceptually motivated by the following factors: (A) perceptual conspicuousness (depending on the use of aggregates); (B) arbitrary divisibility (whether the entity can be divided into portions of any size which are still classified as the original entity, e.g., *machine* vs. *butter*); (C) heterogeneity (whether the entities making a group are of the same or different kind); and (D) how humans interact with the entity (whether they can be seen as individuals or not, e.g., *rice* vs. *pumpkin*).

Additional efforts on noun classification are based on distribution of nouns in corpora and information (cues) from the context to extract information

about the noun (lexical) classes, description and their behaviour.

To test the plausibility of the distributional hypothesis, Hindle (1990) attempts at quasi-semantic classification of nouns observing similarity of nouns based on distribution of subject, verb, object in a corpus. This distributional hypothesis defines reciprocally most similar nouns or reciprocal nearest neighbours – a set of substitutable words, many of which are near synonyms, or closely related.

(Bel et al., 2012) propose a cue-based automatic noun classification in English and Spanish which uses previously known noun lexical classes - event, human, concrete, semiotic, location, and matter. The work is based mainly on (Harris, 1954)'s distributional hypothesis and markedness theory of the Prague Linguistic School, and assumes that lexical semantic classes are properties of a number of words that recurrently co-occur in a number of particular contexts (Bybee, 2010). They use aspects of linguistic contexts where the nouns occur as cues – namely, predicate selectional restrictions (verbal and non-verbal elements such as adjectives and nouns they combine with), grammatical functions, prepositions, suffixes – that represent distributional characteristics of a specific lexical class.

(Bel et al., 2007) work on the acquisition of deep grammatical information for nouns in Spanish using distributional evidence as features and information about all occurrences of a word as a single complex unit. These effort employs 23 linguistic cues for classifying nouns according to an HPSG-based (Head-driven phrase structure grammar) lexical typology (namely the lexicon of an HPSG-based grammars developed in the LKB (Linguistic Knowledge Builder) platform for Spanish). Grammatical features that conform to the cross-classified types are used as they are considered a better level of generalization than the type. These are namely: *mass* and *countable*; plus three additional types for subcategorization: *trans* (nouns with thematic complements introduced by the preposition *de*); *intrans* (noun that has no complements); *pcomp* (where the complements of the noun are introduced by a bound preposition). The combination of features corresponds to the final type.

Our effort as presented here is based on comparison of the semantic primitives of the nouns in

WordNet and the semantic types within the CPA ontology as used in PDEV, in order to outline the directions for further specifying the WordNet semantic classes.

4 CPA ontology

PDEV framework relies on semantic categories called semantic types, which refer to properties shared by a number of nouns that are found in verb pattern (argument) positions. Semantic types are formulated when they have been repeatedly observed in patterns and are organized into a relatively shallow ontology (up to 10 sublevels for some types) – a portion of the ontology – under the type [Liquid] is exemplified on Fig. 1.

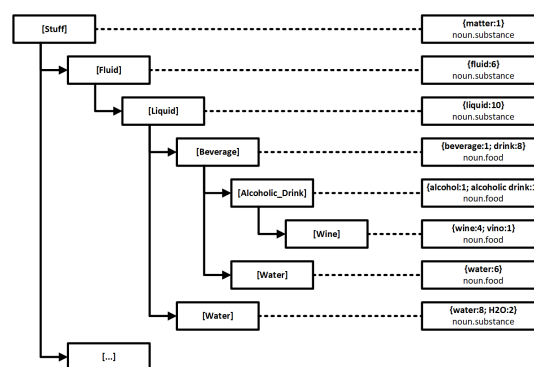


Figure 1: Part of the CPA ontology

On the other hand, some concepts are classified taking into account different properties, such as with drinks – [Beverage] is classified as both [Physical Object] [Inanimate] [Artifact] and [Physical Object] [Inanimate] [Stuff] [Fluid] [Liquid]. As in other ontologies, each semantic type inherits the formal property of the type above it in the hierarchy (Cinkova and Hanks, 2010).

The CPA ontology is language dependent: there are senses of verbs such as *bark* or *saddle* that evoke [Dog] or [Horse] as semantic types because in English there are many words that denote horses and dogs, but there are no verbs that require a distinction between jackals and hyenas, so these are not semantic types (Cinkova and Hanks, 2010).

Though a semantic type usually involves more members than are actually observed in a given pattern position, some words are preferred to others with specific patterns. Therefore, an appropriate level in the ontology should be chosen (the very abstract types such as [Anything] are usually too broad). Thus, the patterns often involve alternative semantic types and not a category, as in the

pattern of the verb *eat*: *[Human] or [Animal] or [Animate] eats ([Physical Object] or [Stuff])*. The alternative larger type can involve types from different levels of the ontology but also can be a type and its supertype. The latter instances are found when a semantic type is predominantly observed in a given pattern position, even if the higher type is also found in the same position.

One of the main indicators of the reliability of semantic types is the fact that they are corpus-driven – they are formulated on the basis of real examples encountered in corpora. Although the semantic types represent cognitive concepts that play a central role in the way words are used, they remain abstract notions as they are not linked to sets of concrete concepts and their lexical representations. Mapping CPA with WordNet will provide sets of concepts and their lexical representations linked to the CPA semantic types.

In addition, in CPA, a single lexical item or a small group of lexical items (called lexical set) that fulfill a role in the clause are included in the verb patterns but not within the ontology (as in: *[Fish] breathes (through gills); [Human] or [Animal] breathes air or dust or gas or [Vapour] (in)*). However, for a precise semantic analysis small sets of lexical items should be represented within the ontology, which implies that the WordNet is the best candidate for full representation of the semantic types ontology.

5 WordNet noun hierarchy

Noun synsets in WordNet are organized into 26 semantic classes (the so-called semantic primitives (Miller et al., 1990)), namely nouns denoting humans (noun.person), animals (noun.animal), plants (noun.plant), acts or actions (noun.act), feelings and emotions (noun.feeling), spatial position (noun.location), foods and drinks (noun.food), etc.

The synsets labeled noun.Tops are the top-level synsets in the hierarchy, the so-called unique beginners for nouns. Thus, the noun synsets are divided into (sub-)hierarchies under the unique noun.Tops labeled synset {entity:1} which has three hyponyms – two unique beginner synsets {physical entity:1} and {abstraction:1; abstract entity:1} and a noun.artifact labeled hyponym {thing:4}. Each of these synsets instantiates a sub-hierarchy. Some of the hyponyms in these sub-hierarchies are also unique beginners. The

hyponyms of the {physical entity:1} synset are:

{thing:1} – noun.Tops containing hyponyms labeled as noun.object;
 {object:1; physical object:1} – noun.Tops, containing hyponyms that are noun.objects and noun.artifacts;
 {causal agent:1; cause:1; causal agency:1} – noun.Tops, containing as hyponyms synsets labeled noun.person, noun.phenomenon, noun.state, noun.object, and noun.substance;
 {matter:1} – noun.substance, containing hyponyms that are noun.substance and noun.object;
 {process:1; physical process:1} – noun.process, with hyponyms marked as noun.process and noun.phenomenon;
 {substance:7} – noun.substance (a sole synset).

Hyponyms of the {abstraction:1; abstract entity:1} synset are (all of these have hyponyms of various semantic class):

{psychological feature:1} – noun.attribute;
 {attribute:1} – noun.attribute;
 {group:1; grouping:1} – noun.group;
 {relation:1} – noun.relation;
 {communication:1} – noun.communication;
 {measure:7; quantity:1; amount:1} – noun.quantity;
 {otherworld:1'} – noun.cognition;
 {set:41} – noun.group.

Though, the basis of classification of certain entities may seem straightforward, it is possible for different entities to inherit information for their features from different (sub-)hierarchies and to have more than one hypernyms, as in (3):

(3)
 {person:1; individual:1; someone:1; somebody:1; mortal:1; soul:1}
 hypernym: {organism:1; being:1}
 hypernym: {causal agent:1; cause:1; causal agency:1}
 (.....)
 hypernym: {physical entity:1}

Additionally, however, there is the EuroWordNet top ontology which contains 63 semantic primitives (Vossen, 1999). The ontology is designed to help the encoding of WordNet se-

semantic relations in a uniform way. The 1st Order Entities are distinguished in terms of main ways of conceptualizing or classifying a concrete entity (Pustejovsky, 1995): Origin, Form, Composition, and Function. Further, Origin is further divided into Natural and Artifact, and Natural – into Living, Plant, Human, Creature, Animal and so on. The 2nd Order Entity is any static situation (property, relation) or dynamic situation, while the 3rd Order Entity is any unobservable proposition which exists independently of time and space (idea, thought).

The WordNet Noun Base Concepts (the most important meanings representing the shared cores of the different WordNets) were classified according to the 1st Order Entity, as follows (Vossen et al., 1998):

(4)

| | |
|--------------------------------|-------------------|
| Artifact | {article:1} |
| Building+Group+Artifact | {establishment:2} |
| Building+Group+Object+Artifact | {factory:1} |

The classification into more than one higher category is a promising approach which is partially followed in our current work.

6 Mapping CPA ontology and WordNet noun hierarchy

We mapped the WordNet noun synset hierarchy onto the semantic type hierarchy in the CPA ontology by matching the CPA semantic types with WordNet synsets and choosing those that are the most probable (and populated) ones, with non-exhaustive results (i.e., many concepts that can be classified under one semantic type, may be not matched under the chosen synsets and left out). Two independent annotators worked on this task and the cases of annotators disagreement were validated by a third one.

Out of 253 instances of matching (one semantic type to one, two, three or more WordNet concepts), there were 46 cases of disagreement between the two annotators; the third annotator worked only on the matches with disagreement, and proposed a new match in 10 instances (in the other cases, the third annotator accepted one of the two choices of the first two annotators; synsets for mapping were selected after an following agreement between the three annotators – in some cases, all suggestions were accepted as matching

options, while in other cases, the annotators agreed on some of the suggestions).

The following general principles were obeyed:

- The WordNet semantic primitives are always preserved.
- New semantic primitives borrowed from the CPA ontology (further called complementary semantic primitives) are supplied added in addition to the WordNet semantic primitives.

To coordinate their work, the annotators agreed on the following:

- The highest appropriate WordNet synset is chosen.
- If necessary, more than one WordNet synset is selected, – in such cases the union of the subtrees is accepted.
- All available PDEV patterns and corpus examples were checked observed to compare them with the WordNet hyponyms belonging to a chosen synset.

As a result of the mapping, the hyponyms of a synset to which a CPA semantic type is mapped, inherit not only the respective WordNet semantic primitive but also the CPA semantic type, as well. For example, all hyponyms of the WordNet synset {location:1} a point or extent in space are classified into with the semantic primitive noun.location. All hyponyms (such as *fact*, *example*, *evidence*, etc.) of the synset {information:2} knowledge acquired through study or experience or instruction mapped with the CPA semantic type [Information] inherit not only the WordNet semantic primitive (noun.cognition) but also the more specific type [Information]. This allows to better prediction for the words connectivity and thus to achieve better results in semantic parsing, word sense disambiguation, language generation and related tasks. The 253 CPA semantic types are mapped to the respective WordNet concepts (synsets) as follows: 199 semantic types are mapped directly to one concept, i.e., [Permission] is mapped to {permission:2} approval to do something, semantic primitive noun.communication; [Dispute] is mapped to {disagreement:2} the speech act of disagreeing or arguing or disputing, semantic prime noun.communication; 39 semantic types are

mapped to two WordNet concepts, i.e., [Route] is mapped to {road:2; route:4;} an open way (generally public) for travel or transportation semantic primitive noun.artifact, and {path:3; route:5; itinerary:3} an established line of travel or access, semantic primitive noun.location; 12 semantic types are mapped to three concepts; 2 semantic types are mapped to four concepts; and 1 semantic type is mapped to five concepts.

Automatic mapping of the hyponym synsets to the inherited CPA semantic types was performed. In the cases where a semantic type and its ancestor were both mapped to the same synset, the ancestor was removed. 82,114 WordNet noun synsets were mapped to the 253 semantic types of the CPA ontology, resulting in 172,991 mappings. As a number of semantic types are classified using different properties, some synsets were mapped to more than one instance of a semantic type, e.g., {phase:6; stage:10} was mapped to both [Abstract_Entity] [Time_Period] and [Abstract_Entity] [Resource] [Asset] [Time_Period]. As these are considered the same concepts, duplicates were removed, leaving 171,359 mappings. The resulting data is available online¹, marked with the XML tag CPA in the WordNet noun synsets.

7 Comparison between WordNet and CPA hierarchies

On the top levels, some classes show a fit between the semantic type and the top level synset, e.g., [Entity] and {entity:1} with subtypes [Abstract_Entity] and {abstract entity:1}, in the most cases the match is not on the same level of the respective hierarchies. For example, [Event] matches {event:1}, but [Event] is on the same level as [Abstract_Entity] in the CPA hierarchy, while {event:1} is linked to the noun.Tops {abstract entity:1} via {psychological feature:1}. Further, [Group] is on the same level as [Entity] but in WordNet {group:1, grouping:1}, which is also noun.Tops, is a hyponym of {abstract entity:1}. Nevertheless, from the fact that not each CPA semantic type can be mapped to one synset, it is clear that the respective nodes in the WordNet hierarchy represent semantic classes and their hyponyms inherit the semantic specifications of the specific semantic class.

If we assume that the concepts are divided into {abstract entity:1} and {physical entity:1} in

WordNet, the types in CPA hierarchy will be marked as follows (we match the CPA subtypes in the respective subhierarchies with probable noun synset(s), which are linked to either of the two noun.Tops; some types below involve subtypes that are matched to WordNet concepts that can be traced back to both {abstract entity:1} and {physical entity:1}) – see on Fig. 2.

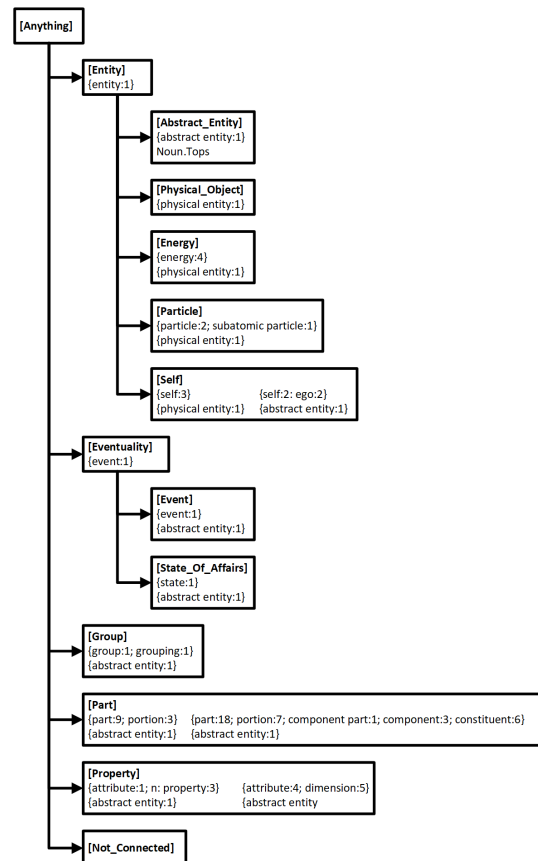


Figure 2: Matching

The matched synsets may be on different levels, and in (5), we exemplify some of the subtypes of the [Artifact] which is a subtype of [Inanimate] under [Physical Object]:

(5)

a. CPA semantic type has two (or more) possible mappings in WordNet, where the synsets belong to different hyponymy paths:

[Artwork]

{artwork:1; art:4; graphics:2; nontextual matter:1} ← {visual communication:1} ← {n: communication:1} ← {abstraction:1; abstract entity:1}

¹http://dcl.bas.bg/PWN_CPA/

{product:2; production:5} ← {n: creation:3} ←
{artifact:1; artefact:1}

[Food]

{food:1; nutrient:1} ← {substance:2} ←
{matter:1} ← {physical entity:1}
{food:3; solid food:1} ← {solid:18} ←
{matter:1} ← {physical entity:1}

b. The WordNet synset to which a CPA semantic type is mapped has two hypernyms:

[Drug]

{drug:3} ← {agent:6} ← {causal agent:1;
cause:1; causal agency:1} ← {substance:2} ←
{physical entity:1}

c. Semantic types that are on the same level in the CPA ontology, are on different levels in WordNet:

[Musical Instrument]

{musical instrument:1; instrument:6} ←
{device:2} ← {instrumentality:1; instrumenta-
tion:3} ← {artifact:1; artefact:1}

[Weapon]

{weapon:1; arm:6; weapon system:1}
← {instrument:5} ← {device:2} ←
{instrumentality:1; instrumentation:3} ←
{artifact:1; artefact:1}

d. Semantic types that are on the same level in the CPA ontology, are direct hypernyms/hyponyms in WordNet i.e., {beverage:1} is a hyponym of {food}

[Beverage]

{beverage:1; drink:8; drinkable:2; potable:2}
← {food:1; nutrient:1} ← {substance:2} ←
{matter:1} ← {physical entity:1}

[Food]

{food:1; nutrient:1} ← {substance:2} ←
{matter:1} ← {physical entity:1}
{food:3; solid food:1} ← {solid:18} ←
{matter:1} ← {physical entity:1}

The following general conclusions can be drawn:

There were certain discrepancies or errors in the

CPA hierarchy as with [Smell] – an attribute – which is included as a subtype of [Vapour] together with [Air] and [Gas] (physical forms of substance); and [Blemish] – again more of an attribute or a result – which is on the same level as [Artifact], [Location], [Structure], [Stuff], etc.

A mismatch was also observed in the hypernym/hyponym structure under the top-level concepts as not every of their hyponyms instantiates another hypernym/hyponym tree (for example {otherworld:1} has no hyponyms, and the notion of cognition is spread throughout both the CPA ontology and WordNet).

New semantic primitives borrowed from the CPA ontology were added to the WordNet structure as complementary semantic primitives and with this the information about co-occurrences between verbs and nouns belonging to particular word classes was enriched and more information encoded/expressed within the WordNet semantic network became explicit.

8 Future work

We plan to automatically assign the PDEV patterns to the WordNet verb synsets and to compare PDEV patterns and WordNet sentence frames. Further, we intend to work on the elaboration of general sentence frames to describe the semantic and syntactic properties of all verb synsets grouped in the verb hypernym/hyponym trees. Testing the semantic compatibility between the general sentence frames and the WordNet semantic primitives (both original and complementary) over corpora examples will help us further elaborate general sentence frames and complementary semantic primitives.

Acknowledgments

The work is done within the project *Towards a Semantic Network Enriched with a Variety of Relations* – DN 10/39/2016, financed by the National Scientific Fund of the Republic of Bulgaria. We would like to thank three anonymous reviewers for their helpful comments, as well as the 9th Global WordNet Conference participants.

References

Anna Korhonen. 2002. Assigning verbs to semantic classes via wordnet. *Proceedings of the 2002 Workshop on Building and Using Semantic Networks*,

- Volume 11, Association for Computational Linguistics.
- Anna Wierzbicka. 1986. What's in a noun? (or: How do nouns differ in meaning from adjectives?) *Studies in Language*, International Journal sponsored by the Foundation Foundations of Language, 10.2 (1986): 353–389.
- Anna Wierzbicka. 1984. Cups and mugs: Lexicography and conceptual analysis. *Australian Journal of Linguistics*, 4.2 (1984): 205–255.
- Anna Wierzbicka. 1985. *Lexicography and Conceptual Analysis*. Karoma, Ann Arbor.
- Anna Wierzbicka. 1988. *The semantics of Grammar*, Volume 18. John Benjamins Publishing.
- Patrick Hanks. 2004. Corpus Pattern Analysis. *Proceedings of Euralex*. Lorient, France.
- Beth Levin. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press.
- Bonnie J. Dorr. 1997. Large-scale dictionary construction for foreign language tutoring and interlingual machine translation. *Machine Translation*, 12(4), 271–322.
- Bonnie J. Dorr and Mari Broman Olsen. 1997. Deriving verbal and compositional lexical aspect for NLP applications. *Proceedings of the Eighth Conference of European Chapter of the Association for Computational Linguistics*, 151–158.
- Christiane Fellbaum, Anne Osehrson, and Peter E. Clark. 2009. Putting semantics into WordNets morphosemantic links. *Proceedings of the Third Language and Technology Conference, Poznan, Poland*, [Repr. in: Responding to Information Society Challenges: New Advances in Human Language Technologies. Springer Lecture Notes in Informatics], vol. 5603, 350–358.
- Donald Hindle. 1990. Noun classification from predicate-argument structures. *Proceedings of the 28th Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics.
- George A. Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine J. Miller. 1990. WordNet: An on-line lexical database. *International Journal of Lexicography*, 3, 4 (Winter 1990), 235–312.
- James Pustejovsky. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.
- Joan Bybee. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- Lei Shi and Rada Mihalcea. 2005. Putting pieces together: Combining FrameNet, VerbNet and WordNet for robust semantic parsing. *Computational linguistics and Intelligent Text Processing*. (2005), 100–111.
- Maria Teresa Pazienza, Marco Pennacchiotti, and Fabio Massimo Zanzotto. 2006. Mixing wordnet, verbnet and propbank for studying verb relations. *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC-2006)*.
- Nuria Bel, Sergio Espeja, and Montserrat Marimon. 2007. Automatic acquisition of grammatical types for nouns. *Human Language Technologies 2007: Proceedings of the North American Chapter of the ACL Companion Volume*, 5–8.
- Nuria Bel, Lauren Romeo, and Muntsa Padr. 2012. Automatic lexical semantic classification of nouns. *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 1448–1455.
- Patrick Hanks and James Pustejovsky. 2005. A Pattern Dictionary for Natural Language Processing. *Revue Française de linguistique applique*, 10:2.
- Patrick Hanks. 2012. Mapping meaning onto use: a Pattern Dictionary of English Verbs. *AACL 2008*, Utah.
- Patrick Hanks. 2013. *Lexical Analysis*. Cambridge, MA: MIT Press.
- Piek Vossen. 1999. EuroWordNet General Document. http://www.globalwordnet.org/ewn/general_document.ps; [October 2017]
- Piek Vossen, Graeme Hirst, Horacio Rodriguez, Salvador Climent, Nicoletta Calzolari, Adriana Roventini, Francesca Bertagna, Antonietta Alonge, and Wim Peters. 1998. *EuroWordNet: A Multilingual Database with Lexical Semantic*. Kluwer Academic Publisher.
- Rebecca Green, Lisa Pearl, Bonnie J. Dorr, and Philip Resnik. 2001. *Lexical Resource Integration across the Syntax-semantics Interface (No. LAMP-TR-069)*. Maryland University College Park Institute for Advanced Computer Studies.
- Silvie Cinkova and Patrick Hanks. 2010. Validation of Corpus Pattern Analysis – Assigning pattern numbers to random verb samples. At: http://ufal.mff.cuni.cz/spr/data/publications/annotation_manual.pdf [October 2017]
- Zellig Harris. 1954. Distributional structure. *Word*, 10(23): 146162.