

# Towards Expanding WordNet with Conceptual Frames

Svetla Koeva

Institute for Bulgarian Language "Prof. Lyubomir Andreychin"

Bulgarian Academy of Sciences

svetla@dcl.bas.bg

## Abstract

The paper presents the project *Semantic Network with a Wide Range of Semantic Relations* and its main achievements. The ultimate objective of the project is to expand Princeton WordNet with conceptual frames that define the syntagmatic relations of verb synsets and the semantic classes of nouns felicitous to combine with particular verbs. At this stage of the work: a) over 5,000 WordNet verb synsets have been supplied with manually evaluated FrameNet semantic frames, b) 253 semantic types have been manually mapped to the appropriate WordNet concepts providing detailed ontological representation of the semantic classes of nouns.

## 1. Introduction

The paper presents and discuss the results of the research project *Semantic Network with a Wide Range of Semantic Relations* (2016 – 2020)<sup>1</sup>. The project targets to expand WordNet with conceptual frames that define the syntagmatic relations of verb synsets and the semantic classes of nouns felicitous to combine with particular verbs.

In Princeton WordNet, each verb synset is associated with a list of sentence frames illustrating the types of simple sentences in which the verbs in the synset can be used (Fellbaum 1990/1993: 55). WordNet sentence frames represent information for the number of frame elements, some semantic information – whether a given element is a human or not, and brief syntactic information – whether the element is realized as a noun, a prepositional phrase (in some cases the preposi-

tion is indicated), an adjective, an *-ing* form of the verb, a clause, an infinitive clause or a *that* clause. For example, the verbs from the synset {*hate; detest*} with a definition ‘dislike intensely; feel antipathy or aversion towards’ are associated with the sentence frames: **Somebody ----s somebody** and **Somebody ----s something**. There are 35 generic frames and a sentence frame might be applicable to all literals within a synset or only to some of them. The frame information given on verbs in WordNet is not sufficient to indicate syntagmatic relations between synsets (syntagmatic relations are semantic relations that express the semantic compatibilities of words). For example, humans and some animals can run, thus most of the nouns from WordNet synsets marked as noun.person and many nouns marked as noun.animal can be linked with the verb *run* as its **Agent**.

To remedy the deficiency of syntagmatic relations in WordNet we introduce the notion of **conceptual frame**, which refers to the set of verbs having equal syntagmatic relations with nouns.

The framework of conceptual frames is built upon the WordNet morphosemantic relations introduced by Miller and Fellbaum (2003). Pre-determined by the meanings of derivational affixes, the morphosemantic links express semantic relations between a verb synset and a noun synset (for example, *an inventor* is an **Agent** of the verb *invent*; *a hanger* is a **Location** of the verb *hang*, a dinner is an **Event** of the verb *dine*, etc.) (Fellbaum et al. 2007). In fact, the morphosemantic relations outline subclasses among the WordNet noun classes: e.g., nouns that can act as human **Agents**, nouns that can act as inanimate **Agents**, etc., and further, the existence of a morphosemantic relation between a verb synset

---

<sup>1</sup> <https://dcl.bas.bg/en/semantichni-mrezhi/>

and a noun synset can serve as an indicator for defining the respective conceptual frame.

The enrichment of the WordNet structure with conceptual frames is related with the implementation of the following steps:

- a) identification of verb synsets that evoke a particular FrameNet semantic frame;
- b) detailed ontological representation of semantic classes of noun synsets;
- c) specification of frame elements relevant for the expression of syntagmatic relations;
- d) assigning the frame elements with noun semantic classes or a combination of classes ensuring the words' compatibility;
- e) definition of WordNet conceptual frames;
- f) insertion of syntagmatic relations within the WordNet structure.

The assumption is that a relatively small number of conceptual frames, which represent the predicate – argument relations between verb and noun synsets, will introduce a large number of syntagmatic relations.

In the presented approach, we take the advantage of automatic mapping of existing resources and rely on the precision of manual assessment of the results. We integrated particular types of semantic knowledge represented basically in three resources: Princeton WordNet 3.0<sup>2</sup> (offering an extensive lexical coverage organized in a semantic network by means of semantic relations), FrameNet<sup>3</sup> (presenting a deep conceptual description of semantic frames), and PDEV (Pattern Dictionary of English Verbs) with the CPA (Corpus Pattern Analysis) semantic types (offering a large ontology of noun semantic classes).

We are going to present here briefly steps a) and b). In particular, we specify the WordNet noun semantic classes into a more fine-grained ontology by mapping WordNet noun hierarchies with the CPA ontology (Section 4) and combine verb hierarchies in WordNet with FrameNet frame semantics and PDEV verb patterns (Section

## 2. Introduction to Conceptual Frames

Conceptual frames are abstract structures which define the semantic and syntactic compatibility between verb predicates and noun arguments. A particular conceptual frame is: associated with a semantic class that expresses its general semantic properties (ideally, each conceptual frame will be assigned with a unique semantic class); represented by a set of verbs organized in the Word-

Net synonym sets, and described by a set of frame elements. The frame verbs can be one or several: linked between each other with lexical relations (synonymy, antonymy) and / or hierarchical relations (hypernymy, troponymy, entailment). The conceptual frame elements roughly correspond to the FrameNet core elements, which means that there is no one-to-one correspondence between FrameNet semantic frames and WordNet conceptual frames.

The selection of conceptual frame elements is based on the intuition about the core participants within a situation but also on the frame elements (implicit or explicit) of superordinates and subordinates in the WordNet hierarchies and on the information from the already available frame representations such as WordNet sentence frames, the FrameNet semantic frames (Ruppenhofer et al. 2016), the VerbNet verb classes (Palmer et al. 2017), the PropBank frames (Bonial et al. 2014), the PDEV patterns (Hanks 2013), the VerbAtlas frames (Di Fabio et al. 2019).

Each conceptual frame element is associated with a set of nouns that are compatible with the verb predicate. Again, the set could contain a single noun or several nouns linked between each other with lexical relations (synonymy, antonymy) and / or hierarchical relations (hypernymy, hyponymy). The association between the frame (verb synsets) and its elements (noun synsets) can be explicitly introduced in WordNet by means of syntagmatic relations. If more than one noun synset can express the frame element (which is the usual case), the syntagmatic relation links the verb synset with the top-most noun synset of the hierarchy, grouping nouns with the same semantic properties (semantic class). The diversity in the compatibilities between representatives of verb classes and noun classes drives the necessity for a detailed ontology of semantic classes.

We can generalize that a **conceptual frame** defines a unique set of syntagmatic relations between: a) verb synsets representing the frame, and b) noun synsets expressing the frame elements (Koeva 2020). Thereby, the notion of conceptual frame combines semantic knowledge presented in WordNet and FrameNet and builds upon it.

The framework of conceptual frames is closely related to the FrameNet semantic frames. Semantic frames are schematic representations of situations involving various participants, props,

---

<sup>2</sup> <http://wordnetweb.princeton.edu> [30 November 2020]

<sup>3</sup> <https://framenet.icsi.berkeley.edu> [30 November 2020]

and other conceptual roles, each of which is a frame element (Johnson and Fillmore 2000: 56). The semantic frames contain frame elements which have a name, a definition, a semantic type, a specification for their core status, and frame internal relations among the frame elements. The main difference between conceptual frames and the FrameNet semantic frames is that conceptual frames are explicitly linked with the noun synsets representing the words with which the verb predicate can be combined (to the extent this is possible due to WordNet structure and content and metaphoric language use).

For example, a conceptual frame which roughly corresponds to the FrameNet semantic frame **Experiencer\_focused\_emotion** is represented by the verb synsets: {dislike} ‘have or feel a dislike or distaste for’; {hate, detest} ‘dislike intensely; feel antipathy or aversion towards’; {like} ‘find enjoyable or agreeable’; {love} ‘have a great affection or liking for’. The conceptual frame elements are **Experiencer** and **Content** (if we keep the names of the FrameNet core elements). The semantic classes of nouns that they could be expressed with are [Human], [Animal], [Physical entity], and [Abstraction] and the combinations are the following:

**Experiencer:** {person, individual} – **Content:** {physical entity} ∪ {abstraction}  
or

**Experiencer:** {animal} – **Content:** {physical entity}.

The syntagmatic links which can be introduced are:

{dislike} and {hate, detest} and {like} and {love} have **Experiencer<sub>1</sub>** {person, individual} and have **Content<sub>1</sub>** {physical entity} and have **Content<sub>1</sub>** {abstraction};

{dislike} and {hate, detest} and {like} and {love} have **Experiencer<sub>2</sub>** {animal} and have **Content<sub>2</sub>** {physical entity}.

One verb synset can be linked by means of one and the same syntagmatic relation with either one or many noun synsets. Many to many syntagmatic relations do not exist.

Ideally, the conceptual frame of the top-most verb in a hierarchy should be the same as the frames of its subordinates. However, it is noticed that troponymy actually comprises various types of manner relation. For example, verbs of motion may specify the kind of transportation (*train, bus, truck, bike*) or the speed dimension (*walk, run*) (Talmy 1985: 62–72; Fellbaum 1990/1993: 47). This implies that verb hierarchies may be elaborated further and verb semantic classes (trees) may be divided in a more precise way. This would result in smaller trees; however, the generalizations for conceptual frames related

with these trees would be more precise. Other problems with the generalizations in conceptual frames might arise from the way of conceptualization (for English or other languages), the level of granularity, the lack of consistency in representing causative and inchoative verbs, the lack of consistency in representing verb aspects for languages expressing this category and some others.

As for conceptual frames (if they are correctly defined), we can expect that the daughter verb synsets will inherit the conceptual frame assigned on the top of the verb tree and deviations are expected in two directions: differences in the explicitness of core frame elements and a reduction of the members of the set of nouns eligible to express a particular frame element (a general tendency is that verbs expressing more specific manners enforce more specific restrictions).

For example, the hyponyms of the verb {dislike} can be linked with the following syntagmatic relations:

{abhor, loathe, abominate, execrate} ‘find repugnant’ has **Experiencer** {person, individual} and has **Content** {physical entity};

{contemn, despise, scorn, disdain} ‘look down on with disdain’ and {look down on} ‘regard with contempt’ have **Experiencer** {person, individual} and have **Content** {person, individual}.

To summarize, some of the main advantages of both resources (WordNet and FrameNet) with regard to the conceptual description of the predicate – argument structure can be complemented and upgraded to expand WordNet with conceptual frames that represent verb predicate – argument syntagmatic relations.

### 3. Combining Semantic Information from Existing Semantic Resources

There are many rich semantic resources (mainly for English but also for other languages) that include different types of semantic information: WordNet (Miller et al. 1990/1993), FrameNet (Baker et al. 1998), VerbNet (Kipper et al. 2008), PropBank (Palmer et al. 2005), Ontonotes (Weischedel et al. 2011), PDEV (Hanks 2004), Yago (Suchanek et al. 2007), BabelNet (Navigli, Ponzetto 2012), VerbAtlas (Di Fabio et al. 2019), SynSemClass (Urešová et al. 2020), among others.

The main advantages of WordNet for semantic analysis focused on introducing conceptual frames are: a) the large number of concepts organized in a semantic network; b) the grouping of concepts in semantic classes according to their general meaning. The main advantages of FrameNet for implementing conceptual frames

are: a) the extensive description of semantic knowledge about an event type and its participants; b) the linking semantic frames with semantic relations. The main advantages of PDEV with CPA for the specification of conceptual frame elements are: a) a description of the semantic types of the elements of verb patterns; b) the organization of semantic types in a shallow ontology. Below we briefly discuss the advantages of the three resources.

### 3.1. Princeton WordNet

WordNet (Miller 1986; Miller et al. 1990/1993: 1–9; Miller, Fellbaum 1991; Fellbaum 1998) is a lexical semantic resource that provides diverse and wide-ranging semantic information. In WordNet, the hypernymy relation (and its inverse relation, hyponymy) links more general concepts to more specific ones and organizes the noun synsets in hierarchies with the most abstract concepts being at the root of trees and the most specific concepts at the leaves of trees (Miller et al. 1990/1993: 12). The hierarchies of verbs are shallow: verbs at the roots of trees express more abstract concepts, while verbs at lower levels of the trees (troponyms) express more specific concepts that denote the manner of doing something (Fellbaum 1990/1993: 47). The inheritance principle of *is-a* relations (such as hypernymy and hyponymy/troponymy) states that anything that is true about the generic entity type A, must also be true about the specific entity type B. Any attributes of A, therefore, are also attributable of B (but not necessarily vice versa). Similarly, in whichever relation A can participate, B can participate also (Storney 1993: 461). In WordNet, a hyponym inherits all the features of the more generic concept and adds at least one feature that distinguishes it from its superordinate and from any other hyponyms of that superordinate (Miller et al. 1990/1993: 8).

Nouns and verbs are grouped in WordNet into more specific semantic classes (Miller 1990/1993: 16; Fellbaum 1990/1993: 41), describing their general meaning: noun.person, noun.animal, noun.cognition; verb.cognition, verb.change, etc. Nouns are classified into twenty-five semantic classes and verbs – into fifteen semantic classes: fourteen classes for events or actions and one class for verbs denoting states (Fellbaum 1990/1993: 41). For example, the verb synonyms {cook; fix; ready; make; prepare} with a definition ‘prepare for eating by applying heat’ have a sentence frame **Somebody ----s something** and a semantic class verb.creation which is inherited by their hyponyms like *dress out, deglaze, scallop, escallop, flambe, devil, precook*, etc. However, not every noun classified

as noun.person can collocate with these verbs as their subject and not every noun that is not classified as noun.person can be their object (*the ex-spouse, ?the neoliberal, \*the infant cooks dinner, ?elephant, \*books*). In other words, the WordNet noun semantic classes could be further specified in order to correlate precisely with the verb-noun selectional preferences. An interdependence between the semantic classes of verbs and the sentence frames applicable to the verbs of one and the same class can be tracked, but such task is very ambiguous because of the small number of semantic classes and the small number of different sentence frames in WordNet. This implies that verb hierarchies may be elaborated further and verb semantic classes may also be divided in a more detailed way.

The following semantic information encoded in WordNet is most important for our research: the relations of inheritance in noun and verb synset trees; the semantic classes to which the noun and verb hierarchies belong; and the sentence frames assigned to the verb synsets. Language independent data can be shared while language specific properties are maintained (Bond et al. 2016).

### 3.2. Berkeley FrameNet

FrameNet is another language resource that contains lexical and conceptual knowledge (Fillmore 1982; Fillmore and Baker 2010; Ruppenhofer et al. 2016). FrameNet can be viewed as a semantic network (or a set of small semantic nets), whose nodes indicate the semantic frames and whose arcs represent semantic relations between frames. For the purposes of the presented research, the following information is employed: the sets of verb lexical units related with semantic frames, the inheritance relation between semantic frames, and the description of core and peripheral frame elements and their semantic types.

In FrameNet, all lexical units evoking a semantic frame have identical (or closely comparable) semantic descriptions: they denote the same part of a scene; have the same number and types of frame elements and the same relations between frame elements (Ruppenhofer et al. 2016: 11). For example, the verb *hate*, together with the verbs *abhor, abominate, adore, delight, despair, despise, detest, dislike, dread, empathize, enjoy, envy, fear, grieve, like, loathe, love, luxuriate, mourn, pity, relish, resent, rue, savour* (and some nouns, adjectives, and adverbs), evokes the frame **Experiencer focused emotion**. One and the same semantic frame might be evoked by lexical units which are encoded either as synonyms, or as hypernyms and hyponyms in the WordNet semantic structure. For example, the

verb *hate* is a synonym of the verb *detest* in a synset expressing the meaning defined as ‘dislike intensely; feel antipathy or aversion towards’. The synset {*hate, detest*} has a hypernym {*dislike*} with a definition ‘have or feel a dislike or distaste for’, a sister synset {*resent*} with a definition ‘feel bitter or indignant about’ and two hyponyms: the synset {*abhor, loathe, abominate, execrate*} with a definition ‘find repugnant’, and the synset {*contemn, despise, scorn, disdain*} with a definition ‘look down on with disdain’. The verbs *loathe, execrate, contemn, scorn, disdain* are presented in WordNet only.

FrameNet includes a network of relations between frames. Several types are defined, of which the most important are: **Inheritance** (an *is-a* relation, the child frame is a subtype of the parent frame), **Using** (the child frame presupposes the parent frame as background); **Subframe** (the child frame is a sub-event of a complex event represented by the parent); **Perspective on** (the child frame provides a particular perspective on an unperspectivized parent frame) (Puppenhofer 2016: 80–83). Inheritance is the strongest relation between frames corresponding to an *is-a* relation in many ontologies. The basic idea of this relation is that each semantic fact about the parent must correspond to an equally specific or more specific fact about the child (Puppenhofer 2016: 80).

FrameNet allows for the characterization of ‘role fillers’ by semantic types of frame elements, which ought to be broadly constant across uses (Ruppenhofer et al. 2016: 12). However, not all frame elements are supplied with a semantic type or the semantic types are too general, and in some cases, they do not show the actual restrictions for lexical combinations. For example, the following frame elements of the semantic frame **Experiencer focused emotion** are equipped with semantic types: **Content** with the semantic type [Content]; **Event** with the semantic type [State of affairs]; **Experiencer** with the semantic type [Sentient]; **Degree** with the semantic type [Degree]; **Explanation** with the semantic type [State of affairs]; **Manner** with the semantic type [Manner]; **Time** with the semantic type [Time]. In summary, the lexical units in FrameNet are not grouped into semantic classes and the semantic types of frame elements, if any, are too general to characterize the class of words that can express the frame element (the annotation part of FrameNet illustrates the specific lexical and grammatical realization of the frame elements).

FrameNet contains extensive semantic information for the semantic frames which are evoked by the sets of lexical units. The value of the semantic information is intensified by the organiza-

tion of the semantic frames in a semantic network.

### 3.3. Pattern Dictionary of English Verbs

The third semantic resource is the Pattern Dictionary of English Verbs (PDEV), where the verb arguments are described by means of the semantic types from the Corpus Pattern Analysis. The verb patterns capture the typical uses of verbs in context and represent the basic ‘argument structure’ of each verb (with semantic values stated for each of the elements of the patterns) (Hanks 2004: 87). The patterns consist of a fixed ordered set of semantic categories (the CPA), whose order corresponds to grammatical categories. The CPA semantic types refer to properties shared by a number of nouns that are found in verb pattern (argument) positions. The reliability of the semantic types is due to the fact that they are corpus-driven – they are formulated on the basis of real examples encountered in corpora.

This semantic resource can also be viewed as a semantic network whose nodes indicate the CPA semantic types and directly point to the subjects, objects, complements, and other positions within the verb patterns. The most important part of this semantic resource is the ontology of semantic types describing the properties of lexical units which are appropriate for filling the slots of verb patterns.

## 4. Ontology of Semantic Classes of Nouns

The semantic classes of nouns and verbs in WordNet might be subdivided into a set of semantic subclasses. For example, within the semantic class [Food] we can introduce the subclass of [Beverage] for nouns associated with verbs like *stir, sip, drink, lap*, etc. Such representation aims to specify the organization of concepts into an ontological structure which allows inheritance between the semantic classes down the hierarchy and ensures more precise specification of verb – noun compatibility.

One potential to extend the repository of WordNet semantic classes is to map the WordNet synsets to an existing hierarchy of semantic types, such as the CPA types. The semantic types (e.g. [Human], [Animal], [Part], etc.) refer to properties which can be expressed by words regularly found to participate in particular verb pattern positions (Hanks 2012: 57–59). In other words, the semantic types state the semantic preferences of verbs that determine the sets of nouns and noun phrases that are normally found in a particular clause role depending on a verb predicate. The CPA semantic types are organized

in a shallow ontology which is based on the analysis of corpus data and which could be supplemented with new semantic types if such appear in new verb patterns. Some verb patterns may contain very general preferences, i.e., the semantic type [Anything], while others impose preferences for a limited set of lexical units grouped into more particular semantic types. For example, some verbs are associated with nouns characterized as [Body part]; however, the verb *shampoo* is associated with a more particular semantic type [Hair], the same is referred to the verb *nod*, which is associated with the type [Head], etc. Some verb patterns require a very small set of lexical units for a particular slot and in this case a semantic type is not formulated; instead, the concrete lexical units are listed in the verb pattern.

The expansion of WordNet semantic classes with CPA semantic types is performed manually by matching the CPA semantic types with WordNet synsets and choosing the most appropriate ones (Koeva et al. 2018).

The following general principles were followed:

- The WordNet semantic classes are preserved. New semantic types borrowed from the CPA ontology are attached to the WordNet synsets.
- The highest appropriate WordNet synset is chosen (within the hypernymy tree).

As a result of the mapping, the hyponyms of a synset to which a semantic type is mapped inherit not only the respective WordNet semantic class, but also the CPA semantic type. For example, the hyponyms of the WordNet synset {medium of exchange; monetary system} ‘anything that is generally accepted as a standard of value and a measure of wealth in a particular country or region’ mapped with the semantic type [Money] (for example, *currency*, *cash*, *paper money*, etc.) inherit not only the WordNet semantic class noun.possession, but also the more specific type [Money].

The 253 CPA semantic types are manually 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} ‘approval to do something’, semantic class noun.communication; [Dispute] is mapped to {disagreement} ‘the speech act of disagreeing or arguing or disputing’, semantic class noun.communication; 39 semantic types are mapped to two WordNet con-

cepts, i.e., [Route] is mapped to {road; route} ‘an open way (generally public) for travel or transportation’, semantic class noun.artefact, and {path; route; itinerary} ‘an established line of travel or access’, semantic class 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 (Koeva et al. 2018).

Automatic mapping of hyponym synsets to the inherited 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<sup>4</sup>. As there are multiple hypernymy relations in WordNet some of the inheritances are not correct, and further, the inheritance by multiple hypernyms will be manually evaluated, and if necessary, adjusted.

## 5. Mapping Verb Frames to WordNet

There are previous efforts at linking WordNet with different semantic resources such as FrameNet, VerbNet, PropBank, Levin’s classes (Korhonen 2002; Shi and Mihalcea 2005; Palmer 2009; Baker and Ruppenhofer 2002; Fellbaum and Baker 2008; Baker and Fellbaum 2009; Fellbaum 2010; Tonelli and Pighin 2009; Laparra, Rigau 2010; Palmer et al. 2014; among others). These efforts resulted in different (but limited) coverage of the mapping and are hardly compatible because they use different release versions of WordNet, FrameNet, VerbNet and PropBank.

In our approach we rely on automatic mapping, automatic prediction for the mapping extension and manual evaluation of the results, something which has not been offered so far. All considered resources are manually crafted and our understanding is that their upgrading and extension (facilitated by automatic methods) should be manually evaluated and proved.

### 5.1. Mapping FrameNet Frames to WordNet

The new WordNet to FrameNet mapping is based on three lexical mappings: 2,817 direct mappings provided within FrameNet (Baker and Fellbaum 2009), 3,134 from eXtendedWordFrameNet (Laparra and Rigau 2010), and 1,833 from MapNet (Tonelli and Pighin 2009), and on 1,335 structural mappings with VerbNet. All in all, the unification of mappings resulted in 4,306 unique mappings of a WordNet synset onto a FrameNet frame (Leseva and Stoyanova 2020).

<sup>4</sup> [http://dcl.bas.bg/PWN\\_CPA/](http://dcl.bas.bg/PWN_CPA/)

The procedures applied to improve and extend mapping coverage are based mainly on the relations of inheritance within WordNet and FrameNet. The frames assigned to 250 out of the 566 root verb synsets were manually evaluated: 75 mappings were corrected and 27 root synsets were additionally assigned a semantic frame. As a general procedure, the hypernym's frame was transferred to its hyponyms in the cases where the hyponyms are not directly mapped to FrameNet frames. As a result, **13,226 synsets** were automatically assigned a FrameNet frame

Further procedures were applied aiming at improving the quality of the mapping: a) checks for unmapped WordNet synsets and FrameNet frames; b) automatic or semi-automatic consistency checks; c) manual evaluation of the assigned frames (Leseva and Stoyanova 2020).

For synsets with frames inherited from their hypernyms, the following tests were applied:

- Searching for an additional match between literals in the given synset and the FrameNet lexical units in the related and sister frames; in any other frame in FrameNet; and in the frames assigned to the synset hyponyms and sisters.
- Calculation of similarity between the gloss of a verb synset and FrameNet lexical unit definitions, as well as between the glosses of derivationally related synsets and their hypernyms and FrameNet lexical unit definitions.
- Searching for a match between literals and words contained in the FrameNet frame name.

As a result of these steps, 9,341 new suggestions of more specific or other possible frames have been made for 5,661 synsets with inherited frames from their hypernyms.

Among all mappings 5,025 frames assigned to verb synsets in WordNet have been manually validated by experts<sup>5</sup>.

Further, some frame elements and their subtypes are analyzed with regard to the selectional preferences imposed on their lexical expression (Leseva et al. 2020). Most of the frame elements are complex structures which prepossess a variety of more specific elements. For example, the frame element [Theme] can be characterized as not having control over the situation and not undergoing changes in its structure, form, function or essential properties; some of the defined subtypes of the [Theme] are: [Effected entity] associated with the synset {entity}; [Suspect] associ-

ated with the synsets {person, individual}  $\cup$  {social group}; [Clothing] associated with the synset {clothing, article of clothing}, and so on. [Effected entity], [Suspect] and [Clothing] (and other sub-types) can be viewed as candidates for enriching the system of WordNet semantic types.

## 5.2. Mapping PDEV Patterns to WordNet

Mapping the PDEV verb patterns and WordNet sentence frames is used for expanding WordNet provided that: a) the semantic types from the CPA ontology are featured as arguments of a given predicate in the PDEV patterns; b) the WordNet noun synset hierarchy is already mapped onto the semantic type hierarchy in the CPA ontology.

A set of translation rules was applied to convert PDEV patterns into WordNet sentence frames and to preserve information of optional pattern arguments and alternative semantic types (Koeva et al. 2019a). After translating the PDEV patterns to WordNet frames, the result was used to assign patterns to the verb synsets in WordNet. For the assignment, we assumed the following:

For a synset  $S$  and a literal  $L \in S$ , PDEV pattern  $P \in \text{patterns}(L)$  can be assigned to  $S$  if and only if  $\text{frames}(S) \cap \text{translations}(P) \neq \emptyset$ .

We automatically assigned 2,904 of 4,048 unique PDEV verb patterns to 2,593 of the 13,767 verb synsets in WordNet by matching the verbs in the PDEV patterns to the literals and the translations of the patterns to the sentence frames of the synsets. This resulted in 6,898 synset pattern assignments (a single pattern may be assigned to more than one synset). 358 unique PDEV verb patterns were assigned to 148 of the 561 top verb synsets (altogether 453 synset pattern assignments).

The automatic mapping was subjected to manual validation (Koeva et al. 2019b). The exact matches were few and covered mainly one place predicates and two place predicates without (or with a few) alternative semantic types. In most cases, WordNet sentence frames were less detailed and involved only the obligatory arguments, while the PDEV patterns involved other constituents (adverbials, optional constituents, etc.), hence, it was expected for WordNet sentence frames to match the PDEV patterns only partially.

In cases where both the WordNet sentence frame and the PDEV pattern were evaluated as correct, but the PDEV pattern contained more information, we took the syntactic and semantic information from the PDEV pattern and the addi-

<sup>5</sup> <https://dcl.bas.bg/en/semantic-relations-data/>

tional CPA semantic types were applied to the WordNet sentence frames.

In fact, it is rather rare for patterns to be automatically assigned to more than two literals in a synset, and if they coincide, it is usually with respect to the type of participants (for example, the verbs *{yelp, yip, yap}* were assigned the patterns [Dog] *yelps*, [Dog] *yaps*), and at most with transitive verbs such as [Human] *watches* [Event], [Human] *sees* [Event]. The effect of manual validation and correction is shown at Table 1.

Total number of WordNet verb synsets covered by PDEV	<b>3,220</b>
<b>Confirmed assignments</b>	
Synsets with fully confirmed pattern assignment	1,488
Confirmed pattern assignments for all synsets	4,084
<b>Manually added assignments</b>	
Synsets to which new patterns were manually assigned	930
Manually assigned patterns in total for all synsets	1,568
<b>Automatic assignments, removed at validation</b>	
Synsets from which patterns were removed	1,143
Removed patterns from all synsets	2,815

Table 1: Manual validation of mapping WordNet sentence frames with PDEV patterns

The manually validated PDEV patterns were added to the XML version of the Princeton WordNet verb synsets used for this study, which is publicly available under the CC by license<sup>6</sup>.

## 6. Conclusion and Future Work

The definition of conceptual frames representing the syntagmatic relations between verb synsets from a particular semantic class and noun synsets from particular semantic classes is (largely) language independent and applicable to any word-

net and other semantic networks. In general, languages differ in the syntactic, morphologic and lexical realization.

At this stage of our work we: a) supplied over 5,000 WordNet verb synsets with manually evaluated FrameNet semantic frames, b) provided a detailed ontological representation of the semantic classes of nouns in WordNet.

Further, selected verb synsets (part of basic vocabulary) will be analyzed with respect to: the FrameNet semantic frames assigned to the verbs with special focus to the core elements; the corresponding sentence frames in WordNet; as well as the PDEV verb patterns assigned to the verb synsets with a particular attention to the CPA semantic types. Through the course of the research other available semantic resources might be analyzed for comparison and evaluation of findings. The study will result in the formulation of conceptual frames represented by a set of verbs and described by a set of frame elements.

As it was pointed out, the main difference between a conceptual frame and a semantic frame (as defined in FrameNet) is that the structure of the conceptual frame includes description of the admissible classes of nouns that may be realized as elements of the frame. Thus, the definition of conceptual frames presupposes the explicit insertion of syntagmatic relations in WordNet and contributes to the effort directed to the enrichment of WordNet structures with multiple relations.

The obtained semantic and syntactic information will be analyzed both through corpus studies of the contexts in which the target verbs occur, as well as through manual evaluation by experts. Where necessary, the conceptual frames will be aligned with the data obtained from the corpus analysis and the conclusions of experts.

The presented research may contribute both to theoretical and contrastive linguistic studies and to the implementation of methods for syntactic parsing and semantic role labelling, important NLP tasks with applications in semantic analysis, word sense disambiguation, language understanding and generation and machine translation.

Conceptual frames offer opportunities for more precise (although still probabilistic) description of syntactic dependencies and semantics of frame elements. The integration of syntagmatic relations in WordNet structure will reveal the existing preferences in word compatibilities.

<sup>6</sup> [http://dcl.bas.bg/PWN\\_PDEV/](http://dcl.bas.bg/PWN_PDEV/)



## Acknowledgments

The studies published in this volume have been carried out as part of the projects: *Towards a Semantic Network Enriched with a Variety of Semantic Relations* funded by the National Science Fund of the Ministry of Education and Science of the Republic of Bulgaria under the Fundamental Scientific Research Program (Grant Agreement 10/3 of 14.12.2016) and *Enriching the Semantic Network WordNet with Conceptual Frames* funded by the National Science Fund of the Ministry of Education and Science of the Republic of Bulgaria under the Fundamental Scientific Research Program (Grant Agreement 49/13 of 30.11.2020).

## References

- Colin Baker, Charles J. Fillmore and John B. Lowe. 1998. The Berkeley FrameNet Project. *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, Montreal, Canada, 10–14 August 1998:86–90.
- Colin Baker and Josef Ruppenhofer. 2002. FrameNet's Frames vs. Levin's Verb Classes. Patrick Chew (ed.) *Proceedings of 28th Annual Meeting of the Berkeley Linguistics Society*:27–38.
- Collin Baker and Christiane Fellbaum. 2009. WordNet and FrameNet as Complementary Resources for Annotation. *Proceedings of the Third Linguistic Annotation Workshop (ACL-IJCNLP'09)*, ACL, Stroudsburg, PA, USA:125–129.
- Francis Bond, Piek Vossen, John McCrae and Christiane Fellbaum. 2016. CILI: The collaborative interlingual index. *Proceedings of the Eighth Global WordNet Conference*. Bucharest:50–57.
- Claire Bonial, Julia Bonn, Kathryn Conger, Jena Hwang and Martha Palmer. 2014. PropBank: Semantics of New Predicate Types. *The 9th edition of the Language Resources and Evaluation Conference*. Reykjavik, Iceland:3013–3019.
- Christiane Fellbaum. 1990/1993. English Verbs as a Semantic Net. *International Journal of Lexicography*, Volume 3, Issue 4, Winter 1990:278–301; reprinted in 1993:40–51.
- Christiane Fellbaum. 1998. (Ed.). *WordNet: An Electronic Lexical Database*. Cambridge (MA): MIT Press.
- Christiane Fellbaum. 2010. Harmonizing WordNet and FrameNet. *IceTAL*:2–3.
- Christiane Fellbaum, Anne Osherson, and Peter E. Clark. 2007. Putting Semantics into WordNet's "Morphosemantic" Links. *Responding to Information Society Challenges: New Advances in Human Language Technologies*. Springer Lecture Notes in Informatics, vol. 5603:350–358. 2008.
- Christiane Fellbaum and Colin Baker. 2008. Can WordNet and FrameNet be Made "Interoperable"? Jonathan Webster, Nancy Ide, and Alex Chengyu Fang. (Eds.) *Proceedings of the 1st International Conference on Global Interoperability for Language Resources*, Hong Kong, City University: 67–74.
- Charles J. Fillmore. 1982. Frame semantics. *Linguistics in the Morning Calm*. Seoul: Hanshin:111–137.
- Charles J. Fillmore and Collin Baker 2010. A frames approach to semantic analysis. *Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press:313–341.
- Andrea Di Fabio, Simone Conia and Roberto Navigli. 2019. VerbAtlas: a Novel LargeScale Verbal Semantic Resource and its Application to Semantic Role Labelling. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, Hong Kong, China, November 3–7:627–637.
- Peter Hanks. 2004. Corpus pattern analysis. *Proceedings of the 11th EURALEX International Congress, (EURALEX 2004)*, Lorient, France, 6–10 July 2004:87–97.
- Peter Hanks. 2012. How people use words to make meanings: Semantic types meet valencies. Alex Boulton and James Thomas (Eds.). *Input, Process and Product: Developments in Teaching and Language Corpora*. Brno: Masaryk University Press: 54–70.
- Peter Hanks. 2013. *Lexical Analysis*. Cambridge (MA): MIT Press.
- Christopher Johnson and Charles J. Fillmore. 2000. The FrameNet tagset for frame semantic and syntactic coding of predicate-argument structure. *Proceedings of the Applied Natural Language Processing Conference (ANLP 2000)*:56–62.
- Karin Kipper, Anna Korhonen, Neville Ryant and Marta Palmer. 2008. A Largescale Classification of English Verbs. *Language Resources and Evaluation*, 42(1):21–40.
- Svetla Koeva. 2020. Semantic Relations and Conceptual Frames. Koeva, S. (ed.) *Towards a Semantic Network Enriched with a Variety of Semantic Relations*, Professor Marin Drinov Publishing House of BAS:7–20.
- Svetla Koeva, Tsvetana Dimitrova, Valentina Stefanova and Dimitar Hristov. 2018. Mapping WordNet Concepts with CPA Ontology. *Proceedings of the 9th Global WordNet Conference (GWC'2018)*. Global WordNet Association, Singapore:70–77.
- Svetla Koeva, Dimitar Hristov, Tsvetana Dimitrova, and Valentina Stefanova. 2019a. Enriching Wordnet with Frame Semantics. *Proceedings of the International Annual Conference of the Institute for Bulgarian Language Prof. Lyubomir Andreychin*

- (Sofia, 14th – 15th May 2019). Vanya Micheva, Diana Blagoeva, Siya Kolkovska, Tatyana Aleksandrova, Hristina Deykova (Eds.). Sofia: Prof. Marin Drinov Press of the Bulgarian Academy of Sciences:300–308.
- Svetla Koeva, Tsvetana Dimitrova, Valentina Stefanova and Dimitar Hristov. 2019b. Towards Conceptual Frames. *Foreign Language Teaching*, 46, 6:551–564.
- Anna Korhonen. 2002. Assigning Verbs to Semantic Classes via WordNet. *Proceedings of the 2002 Workshop on Building and Using Semantic Networks*, vol. 11:1–7.
- Egoitz Laparra and German Rigau. 2010. eXtended WordFrameNet. *Proceedings of the Seventh International Conference on Language Resources and Evaluation*, 17–3 May 2010, Valletta, Malta: 1214–1219.
- Svetlozara Leseva and Ivelina Stoyanova. 2020. Beyond Lexical and Semantic Resources: Linking WordNet with FrameNet and Enhancing Synsets with Conceptual Frames. Svetla Koeva (Ed.). *Towards a Semantic Network Enriched with a Variety of Semantic Relations*. Sofia: Professor Marin Drinov Publishing House of BAS:21–48.
- Svetlozara Leseva, Ivelina Stoyanova, Maria Todorova, and Hristina Kukova. Putting Pieces Together: Predicate-Argument Relations and Selectional Preferences. Svetla Koeva. (Ed.). *Towards a Semantic Network Enriched with a Variety of Semantic Relations*. Sofia: Professor Marin Drinov Publishing House of BAS:49–86.
- George Miller. 1986. Dictionaries in the Mind. *Language and Cognitive Processes*, 1:171–185.
- George Miller, R. Beckwith, Christiane Fellbaum, D. Gross, and K. J. Miller. 1990/1993. *Introduction to WordNet: An On-Line Lexical Database*. *International Journal of Lexicography*, 1990, 3(4), pp. 235–244; reprinted in 1993:1–9.
- George Miller and Christiane Fellbaum. 1991. Semantic networks of English. *Cognition*. 41 (1-3): 197–229.
- George Miller and Christiane Fellbaum. 2003. Morphosemantic links in Wordnet. *Traitement automatique des langues*, 44.2:69–80.
- Roberto Navigli, Simone Ponzetto. 2012. BabelNet: The Automatic Construction, Evaluation and Application of a WideCoverage Multilingual Semantic Network. *Artificial Intelligence*, 193, Elsevier: 217–250.
- Marta Palmer. 2009. Semlink: Linking PropBank, VerbNet and FrameNet. *Proceedings of the 5th International Conference on Generative Approaches to the Lexicon*, Pisa, Italy, 2009:9–15.
- Marta Palmer, Daniel Gildea, Paul Kingsbury. 2005. The Proposition Bank: A Corpus Annotated with Semantic Roles. *Computational Linguistics Journal*, 31:1:71–106.
- Marta Palmer, Claire Bonial, and Diana McCarthy. 2014. SemLink+: FrameNet, VerbNet and Event Ontologies. *Proceedings of Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929–2014)*, Baltimore, USA, ACL:13–17.
- Martha Palmer, Claire Bonial, Claire, and Jena D. Hwang. 2017. VerbNet: Capturing English verb behavior, meaning and usage. (Ed.) Susan Chipman. *The Oxford Handbook of Cognitive Science*. Oxford University Press.
- Josef Ruppenhofer, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson, Collin F. Baker, and Jan Scheffczyk. 2016. *FrameNet II: Extended Theory and Practice*. <[https://framenet.icsi.berkeley.edu/fndrupal/the\\_book](https://framenet.icsi.berkeley.edu/fndrupal/the_book)> [30.11.2020]
- Lei Shi and Rada Mihalcea. 2005. Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robust Semantic Parsing. Alexander Gelbukh, (Ed.). *Computational Linguistics and Intelligent Text Processing*. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, vol. 3406:100–111.
- Veda Storey. 1993. Understanding Semantic Relationships. *The International Journal on Very Large Data Bases*, 2:455–488.
- Fabian Suchanek, Gjergji Kasneci, and Gerhard Weikum. 2007. Yago – A Core of Semantic Knowledge. *Proceedings of the 16th International Conference on World Wide Web*, 16th International Conference on World Wide Web, Banff, Alberta, Canada, 8–12 May 2007:697–706.
- Leonard Talmy. 1985. Lexicalization Patterns: Semantic Structure in Lexical Forms. Timothy Shopen (Ed.). *Language Typology and Syntactic Description*, vol. 3. Cambridge: Cambridge University Press:57–149.
- Sara Tonelli and Daniele Pighin 2009. New Features for FrameNet – WordNet Mapping. *Proceedings of the Thirteenth Conference on Computational Natural Language Learning*, Boulder, USA:455–488.
- Zdenka Urešová, Eva Fucíková, Eva Hajicová, and Jan Hajic. 2020. SynSemClass Linked Lexicon: Mapping Synonymy between Languages. *Proceedings of the Globalex Workshop on Linked Lexicography, Language Resources and Evaluation Conference (LREC 2020)*, Marseille, 11–16 May 2020:10–19.
- Ralph Weischedel, Eduard Hovy, Mitchell Marcus, Marta Palmer, Robert Belvin, Sameer Pradhan, Lance Ramshaw, and Nianwen Xue. 2011. Ontonotes: A large training corpus for enhanced processing. Olive, J., C. Caitlin, J. McCary (Eds.). *Handbook of Natural Language Processing and Machine Translation*. Springer:54–63.