BULGARIAN ACADEMY OF SCIENCES
INSTITUTE FOR BULGARIAN LANGUAGE
PROF. LYUBOMIR ANDREYCHIN

TOWARDS A SEMANTIC NETWORK ENRICHED WITH A VARIETY OF SEMANTIC RELATIONS

Sofia 2020

TOWARDS A SEMANTIC NETWORK ENRICHED WITH A VARIETY OF SEMANTIC RELATIONS

BULGARIAN ACADEMY OF SCIENCES
INSTITUTE FOR BULGARIAN LANGUAGE PROF. LYUBOMIR ANDREYCHIN

TOWARDS A SEMANTIC NETWORK ENRICHED WITH A VARIETY OF SEMANTIC RELATIONS

Edited by Svetla Koeva

Sofia • 2020

Prof. Marin Drinov Academic Publishing House of
Bulgarian Academy of Sciences

Acknowledgments

The studies published in this volume have been carried out as part of the project **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 Programme (Grant Agreement 10/3 of 14.12.2016).

NATIONAL SCIENCE FUND

Editorial address:
Institute for Bulgarian Language Prof. Lyubomir Andreychin,
Bulgarian Academy of Sciences
52 Shipchenski Prohod Blvd., bl. 17
Sofia 1113, Bulgaria
+359 2/ 872 23 02

ISBN 978-619-245-057-1

CONTENTS

Preface

SEMANTIC RELATIONS AND CONCEPTUAL FRAMES

Svetla Koeva

Institute for Bulgarian Language *Prof. Lyubomir Andreychin*
Bulgarian Academy of Sciences

1. Introduction

There are many rich semantic resources (mainly for English but also for other languages) that include different types of semantic information: WordNet (Miller 1995), 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 objectives of our research presented in the book *Towards a Semantic Network Enriched with a Variety of Relations* are to expand the WordNet structure with an extended network of conceptual frames that represent verb predicate – argument semantic relations in a generalised way. Our assumption is that a relatively small number of conceptual frames which define predicate – argument relation pairs (classified with respect to an ontology of semantic classes) introduce a large number of semantic relations.

Conceptual frames are abstract structures, each of which is described by a unique set of semantic relations between: a) the frame (represented by a set of verbs organised in the WordNet synonym sets) associated with a particular semantic class that expresses the semantic properties of the frame; b) frame elements (represented by a set of nouns organised in the WordNet synonym sets) linked with particular semantic classes that express the semantic properties of the frame elements; c) the semantic relations between the conceptual frame and its frame elements (the information about the type of semantic relations is not necessarily explicit as it can be extracted by pairing verb semantic classes and noun semantic classes assigned to particular frame elements).

The definition of conceptual frames is presupposed by:

a) The detailed ontological representation of semantic classes of the noun and verb synonym sets (each semantic class combines synonym sets expressing equivalent semantic properties, for example, *human*, *sentient*, *motion*, *communication*, etc.);
b) The inheritance of the semantic properties between the synonym sets organised in the WordNet hierarchies.

The book *Towards a Semantic Network Enriched with a Variety of Relations* presents the integration of a particular type of semantic knowledge defined in three semantic resources: WordNet[1] (an extensive lexical coverage organised in a semantic network by means of semantic relations), FrameNet[2] (a deep conceptual description of semantic frames' properties), and PDEV (Pattern Dictionary of English Verbs) with the CPA[3] (Corpus Pattern Analysis) semantic types (a large ontology of noun semantic classes). In particular, the conceptual frames combine verb hierarchies in WordNet with FrameNet frame semantics, and specify the WordNet noun semantic classes into a more fine-grained ontology by mapping WordNet noun hierarchies with the CPA ontology. The presented study is focused on verb conceptual frames.

2. Semantic Resources

The lexical-semantic network WordNet (Miller et al. 1990/1993: 1 – 9; Fellbaum 1998) encodes human knowledge about synonyms – words (or multiword expressions) denoting one and the same concept, and the semantic relations between the concepts. The synonym sets (synsets) constitute the nodes of the semantic network which are interconnected by arcs representing the semantic relations. There are two types of relations encoded in the Princeton WordNet: relations between literals (lexical relations such as synonymy, antonymy, and derivation) and relations between synsets (conceptual relations such as hypernymy, hyponymy, holonymy, meronymy, etc.) (Miller et al. 1990/1993: 3 – 8). A weak definition for synonymy makes synonymy relative to a context: *two expressions are synonymous in a linguistic context C if the substitution of one for the other in C does not alter the truth value* (Miller et al. 1990/1993: 6). The definition implies that the WordNet synonyms are cognitive (or propositional) synonyms (Cruse 1986: 290): words used with the same meaning in a given context in which they are interchangeable.

WordNet is a lexical-semantic resource that provides diverse and wide-ranging information, the following parts of which are most important for our research: the relations of inheritance encoded 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. The WordNet semantic network of interconnected concepts is to a great extent language independent and, therefore, allows the creation of interconnected networks for different languages, including for Bulgarian[4] (Koeva et al. 2004). Furthermore, the enrichment of one WordNet with conceptual frames and semantic relations is transferable to other wordnets through alignment with the Princeton WordNet.

FrameNet is another language resource that contains lexical and conceptual knowledge (Fillmore 1982; Fillmore, Baker 2010; Ruppenhofer et al., 2016). FrameNet represents lexical units (a pairing of a word, a multiword expression, or

---

[1] http://wordnetweb.princeton.edu [30 May 2020]
[2] https://framenet.icsi.berkeley.edu [30 May 2020]
[3] http://pdev.org.uk [30 May 2020]
[4] http://dcl.bas.bg/bulnet/ [30 May 2020]

an idiomatic phrase with its meaning) by connecting them with abstract semantic structures (semantic frames). The different word meanings are connected with different semantic frames. Semantic frames are schematic representations of situations involving various participants, props, and other conceptual roles, each of which is a frame element (Johnson, Fillmore 2000: 56). The semantic frames are provided with a name and a definition and contain frame elements which have a name, a definition, a semantic type, a specification for their core status, and frame-internal relations between frame elements. FrameNet can also 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 description of core and peripheral frame elements and their semantic types, the sets of verb lexical units related with semantic frames, and the inheritance relation between semantic frames.

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 (CPA). The verb patterns capture the typical uses of verbs in a 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 whose order corresponds to grammatical categories. Each pattern is linked to an implicature explaining the 'meaning' of the pattern using the CPA semantic types (Hanks 2004: 88). The CPA constitutes a shallow ontology of semantic types. 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.

The main advantages of WordNet for semantic analysis focused on introducing conceptual frames are: a) the large number of concepts organised 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. And the main advantages of PDEV with CPA for the specification of the conceptual frame elements are: a) the description of semantic types of the elements of verb patterns; b) the organisation of semantic types in a shallow ontology.

3. Frame Semantics incorporated in the Semantic Resources

In WordNet, each verb synset is associated with a list of sentence frames which illustrate the types of simple sentences in which the verbs in the synset can be used (Fellbaum 1990/1993: 55). The WordNet sentence frames represent information for the number of frame elements. Frame elements are ordered (subject, object, indirect object, clause) and are supplied with shallow semantic information – whether the element is a human or not, and brief syntactic information – whether the element is realised as a noun, a prepositional phrase (in some cases the preposition is indicated), an adjective, an *ing* form of the verb, a clause, an infinitive clause or a *that* clause. For example, the verbs part of 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[5] and a sentence frame might be applicable to all literals within a synset or only to some of them.

Verbs and nouns 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: {act, action, activity}, {animal, fauna}, {artifact}, {attribute, property}, {body, corpus}, {cognition, knowledge}, {communication}, {event, happening}, {feeling, emotion}, {food}, {group, collection}, {location, place}, {motive}, {natural object}, {natural phenomenon}, {person, human being}, {plant, flora}, {possession}, {process}, {quantity, amount}, {relation}, {shape}, {state, condition}, {substance}, {time} (Miller 1990/1993: 16). Verbs are classified into fifteen semantic classes: fourteen classes for events or actions (verbs of bodily care and functions, change, cognition, communications, competition, consumption, contact, creation, emotion, motion, perception, possession, social interaction, and weather verbs) and one class for verbs denoting states (Fellbaum 1990/1993: 41). The verb synset {*hate*; *detest*} is classified as a verb.emotion.

The lexical units sharing one and the same syntactic frame can be extracted; however, the WordNet sentence frames are too general to serve the goal for finding sets of predicates sharing one and the same semantic behaviour. Furthermore, there is a dependency between semantic classes of verbs and sentence frames applicable to the verbs of one and the same class, but this relation is very ambiguous because of the small number of semantic classes and the small number of sentence frames in WordNet.

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 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**.

The frame elements are divided into core, peripheral, and extra-thematic. A core frame element instantiates a conceptually mandatory component of a frame, while making the frame unique and different from other frames (Ruppenhofer et al. 2016:

[5] https://wordnet.princeton.edu/documentation/wninput5wn [30 May 2020]

23). Core frame elements of the frame **Experiencer_focused_emotion**[6] are: **Content** (what the Experiencer's feelings or experiences are directed towards or based upon); **Event** (the occasion or happening that Experiencers in a certain emotional state participate in); **Experiencer** (experiences the emotion or other internal state); **Topic** (the area about which the Experiencer has the particular experience). The core unexpressed elements are: **Expressor** (it marks expressions that indicate a body part, gesture, or other expression of the Experiencer that reflect his or her emotional state); **State** (the abstract noun that describes a more lasting experience by the Experiencer). The non-core elements are: **Circumstances** (it encodes the circumstances or conditions under which the Experiencer experiences the emotion); **Degree** (the Degree to which the experience occurs); **Explanation** (the explanation for why an Experiencer experiences the particular emotion); **Manner** (the Manner of performing an action); **Parameter** (a domain in which the Experiencer experiences the Content); **Time** (the time at which the Experiencer is in the specified emotional state). Ideally, the number of the core frame elements coincides with the number of elements in the WordNet sentence frames and PDEV patterns. Peripheral frame elements do not uniquely characterise a frame and can be instantiated in any semantically appropriate frame. Extra-thematic frame elements situate an event against a backdrop of another state of affairs or a larger frame within which the reported state of affairs is embedded (Ruppenhofer et al. 2016: 24).

Frame elements are related to the frame and required by it, as well as interrelated directly in a number of ways (Ruppenhofer et al. 2016: 25): some groups of core frame elements act like sets in that the presence of any member of the set is sufficient to satisfy a semantic valence of the predicate; in some cases, the occurrence of one core frame element requires that another core frame element occurs as well; and in some cases, if one of the frame elements in a group of conceptually related frame elements shows up, no other frame element from that group can. In FrameNet, for the frame **Experiencer_focused_emotion** two core sets are defined: {Content, Topic} and {Event, Experiencer, Expressor, State}.

Furthermore, FrameNet allows the characterisation 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, 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, FrameNet contains extensive semantic information for the semantic frames which are evoked by the sets of lexical units. On the other hand, the lexical units are not grouped into semantic classes and the semantic types of frame

[6] https://framenet.icsi.berkeley.edu/fndrupal/frameIndex [30 May 2020]

elements, if any, are too general to characterise the class of words that can express the frame element (the annotation part of FrameNet illustrates the specific lexical and grammatical realisation of the frame elements). 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.

The underlying assumptions for developing PDEV are that there are norms (prototypical uses of words for which patterns can be identified within a corpus) and exploitations (deviations from norms for which patterns can be also identified) (Hanks 2013). A PDEV verb entry consists of a list of numbered patterns (frames) linked to implicatures – explanations of the meaning of the patterns (Hanks 2004: 88). Clause roles in verb patterns are (Hanks 2013: 94 – 95): Subject (the semantic subject of the clause); Predicator (the verb, together with its auxiliaries, if any); Object (direct or indirect); Complement (a phrase that is coreferential either with the subject of the sentence or with the direct object); Adverbial (a prepositional phrase, a particle, or one of a small set of adverbs). The pattern is ordered, with a subject on the left of the verb, and objects, complements and adverbials on its right. The usual fillers of clause roles are nouns that share some aspects of their meaning, which is described as a semantic type (Cinkova, Hanks 2010: 4; Hanks 2012: 66). The semantic types (e.g. [Human], [Animal], [Part], etc.) refer to properties which can be expressed by words regularly found to participate in particular pattern positions (Hanks 2012: 57 – 59). The semantic types are organised in a shallow ontology which is driven by the analysis of corpus data and which could be supplemented with new semantic types if such appear in new verb patterns. Some patterns contain very general preferences, i.e. the semantic type [Anything], while some other distinguish preferences for a smaller set of lexical items grouped into semantic types. For example, some verbs are associated with noun sets characterised 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 take only a very small set of lexical units as normal collocates in a particular slot and in this case a semantic type is not formulated; rather, the lexical units are listed in the verb pattern. In other words, the purpose of the semantic types is to state the semantic preferences that determine the range of nouns and noun phrases that are normally found in a particular clause role. Any pattern position may contain alternative semantic types. For example, the patterns of the verb *hate*, with their implicatures, are defined in PDEV as follows:

1. Human | Animal hate Entity | Group | Part | Property

   [[Human | Animal]] strongly dislikes [[Entity | Group | Part | Property]]

2. Animal | Human hate Activity to+INF ING

   [[Human | Animal]] strongly dislikes or regrets {{to/INF [V] | -ING} | [[Activity]]}

3. Human would hate to+INF

   [[Human]] would be unhappy or disappointed {to/INF [V]}

4. Human hate it THAT WH+

   [[Human]] is very unhappy {it | that-CLAUSE | WH-CLAUSE}

The semantic types [Human] and [Animal] with which the semantic subject is described correspond to the FrameNet frame element **Experiencer** encoded in the semantic frame **Experiencer_focused_emotion.** The semantic type [Human] can be linked with the WordNet verb synset {person; individual; someone; somebody; mortal; soul} and the semantic type [Animal] – with the verb synset {animal; animate being; beast; brute; creature; fauna} each of which dominates a set of hyponyms appropriate to collocate with the verb *hate*.

The PDEV pattern elements correspond to the FrameNet core frame elements. If this correspondence is formally established, each frame element will be supplied with a CPA semantic type. On the other hand, if the correspondence between the CPA semantic types and the WordNet noun synsets is identified, each frame element will be provided with the information for the set of nouns that could express it.

3. Semantic Relations

The relations in the semantic network WordNet are binary. One synset can be linked by means of one and the same relation with either one or many synset nodes. For example, a hyponym can have exactly one hypernym, but a hypernym can have either one or many hyponyms. Many to many relations are not included. Respectively, in FrameNet, the one to one or one to many relations occur between the semantic frames and in CPA, the one to one or one to many relations occur between the semantic types and the elements of verb patterns.

3.1. Inheritance Relations

In WordNet, the hypernymy relation (and its inverse relation, hyponymy: subordination, class inclusion) links more general concepts to more specific ones and organises the noun synsets in hierarchies with the most abstract concepts being at the root of the trees and most specific concepts at the leaves of the trees (Miller et al. 1990/1993: 12). Hypernymy and hyponymy relations satisfy properties for asymmetry and transitivity (Lyons 1977: 291 – 292; Miller, 1990/1993: 13). The asymmetry states that for two terms $a, b \in E$, if $a < b$ then $b \nless a$. The transitivity states that for three terms $a, b, c \in E$, if $a < b$ and $b < c$ then $a < c$. For example, if *bird* is a hypernym of *parrot*, then *parrot* is not a hypernym of *bird*, and vice versa if *parrot* is a hyponym of *bird*, then *bird* is not a hyponym of *parrot*. Another example illustrates the transitivity: if *bird* is a hypernym of *parrot* and *parrot* is a hypernym of *cockatoo*, then *bird* is a hypernym of *cockatoo*. And vice versa, if *cockatoo* is a hyponym of *parrot* and *parrot* is a hyponym of *bird*, then *cockatoo* is a hyponym of *bird*.

The relations of inclusion link a more general entity with a more specific entity. Class inclusion may be expressed with the definitions: X's are a type of Y, X's are Y's; X is a kind of Y, and X is-a Y (respective examples are *Cars are a type of Vehicle*; *Roses are Flowers*; *Theft is a kind of Crime*; and *Employee is-a Person* (Winston et al. 1987: 427). Several types of inclusion are mentioned (Storey 1993: 460): classification, which relates an entity occurrence to an entity type; generalisation, in which an entity type is the union of non-overlapping subtypes; specialisation, which is defined as the inverse of generalisation; and subset hierarchy, in which possibly overlapping subtypes exist. In another description, inclusion is divided into two main types: generic to generic relations (subset and superset, generalisation and specialisation, a kind of, conceptual containment, role value restrictions, set and its characteristic types) and generic to individual relations (set membership, predication, conceptual containment, abstraction) (Brachman 1983: 32). Such analysis implies that the class inclusion hierarchy of noun synsets might be divided into different hierarchies according to the specific type of the inclusion.

Common practice in wordnets is to use multiple hypernyms. Multiple hypernyms can be exclusive (*albino is either an animal or a human*), conjunctive (*spoon is both cutlery and container*) or non-exclusive (*knife can be cutlery, a weapon, or both*) (EAGLES 1999: 212). Based on the assumption that one synonym set cannot be related with more than one hypernym, other semantic relations are defined in the scope of multiple hypernymy: origin, form and function, as well as the relation true hypernymy (Koeva et al. 2018: 368 – 371). Multiple hypernymy embraces several semantic relations which, in turn, are not labelled and are only partially shown within the WordNet structure. This means that hypernymy (the general inheritance relation) might be further specified and divided into several relations. Such specification would better outline the subsets of nouns that saturate semantic preferences of a verb predicate within the semantic classes of nouns, which are propagated through the inheritance (hypernymy) relation.

The verb synsets are arranged into hierarchies which express entailment relations (Fellbaum 1990/1993; Fellbaum 2002). In Wordnet, four types of entailment are recognised (Fellbaum 1993: 54): troponymy, which represents a special case of entailment – pairs that are always temporally co-extensive like *limp* and *walk, lisp* and *talk* (Fellbaum 1990/1993: 47); proper temporal inclusion is another type of entailment, in some wordnets it is described as a subevent relation between pairs like *snore* and *sleep, buy* and *pay* (Piasecky, Koeva 2017: 292); backward presupposition is an entailment relation without temporal inclusion of pairs like *succeed* and *try, untie* and *tie*; cause is an entailment relation without temporal inclusion of pairs like *give* and *have, expel* and *leave*.

The hierarchies of verbs are shallow: verbs at the roots of the 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, 2002). Within a given semantic class of verbs there are several independent trees (Fellbaum 1990/1993: 48) which could be artificially joined to an abstract top synset. The semantic class of verbs determines the specific manner that is expressed, and verbs placed towards the bottom of the trees (troponyms) convey increasingly specific manners. It is also 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). Thus, subsets of particular kinds of manners tend to cluster within a given semantic field (Fellbaum 1990/1993: 47). This implies that verb hierarchies may be elaborated further and verb semantic classes may be divided in a more precise way. This would result in smaller trees; however, the generalisations for conceptual frames related with these trees would be more precise.

The inheritance principle of is-a relations states that anything that is true about the generic entity type A, must also be true of the specific entity type B. Any attributes of A, therefore, are also attributable to 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). It is stated also that, as one descends in a verb hierarchy, the variety of nouns that the verbs on a given level can take as potential arguments decreases, which seems to be a function of the increasing elaboration and meaning specificity of the verb (*walk* can take a subject referring either to a person or an animal; most troponyms of *walk*, however, are restricted to human subjects) (Fellbaum 1990/1993: 49).

Inheritance is important in the way that all noun synsets that are hyponyms of a synset representing a particular semantic class should inherit the properties of this class, and also all verb synsets that are hyponyms of a synset associated with a particular conceptual frame should inherit the properties of this frame. This is generally true and, if the inheritance relations of nouns and verbs are further specified, noun synset hierarchies can serve as sets of words eligible to fill in particular verb predicate slots, while verb synset hierarchies can be organised according to a general conceptual frame.

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 subevent of a complex event represented by the parent); Perspective on (the child frame provides a particular perspective on an un-perspectivised parent frame) (Puppenhofer 2016: 80 – 83). Inheritance is the strongest relation between frames corresponding to 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). The following complications were observed: a daughter frame may have frame elements not listed in the parent frame, or these may be extra-thematic in the parent frame; a daughter frame often does not mention frame elements of the parent that have the type "Core-unexpressed"; two frame elements of a parent may map onto one frame element of the daughter (Puppenhofer 2016: 81 – 82).

As for the conceptual frames (if they are correctly defined) within a fine-grained WordNet structure of inheritance relations we can expect that the daughter verb synsets will inherit the conceptual frame assigned on the top of the tree and deviations are expected in two directions: a reduction of a core frame element and a reduction of the members of the set of nouns eligible to express a particular frame element.

3.2. Case Relations

The relations involved in predication or attribution have been described as case relations (Fillmore 1968). A taxonomy of seven types of semantic relations (inclusion, possession, attachment, attribution, antonym, synonym, and case) is presented (Storey 1983: 456) based on the work of Winston et al. (1987), Chaffin et al. (1988) and Landis et al. (1987). Most important for establishing the network of conceptual frames are the semantic relations of inclusion, synonymy and case. Five different types of case relation were identified by Chaffin and Herrmann (1984: 136 – 137). Three involve agents (agent-action, agent-instrument, and agent-object), and two involve actions (action-recipient and action-instrument) (Storey 1983: 470 – 471). For example, a typical agent for the activity of *barking* is a *dog*; the agent-action relation is exemplified by pairs such as *dog-bark* (Chaffin, Herrmann 1984: 136).

Miller and Fellbaum (2003) describe the addition to WordNet of morphosemantic relations that connect derivationally related words. Fellbaum et al. (2007) suggest that the meanings of derivational affixes can be classified into a relatively small number of semantic categories (agent, instrument, etc.), which actually represent the semantic nature of morphosemantic relations. A morphosemantic relation is considered as a kind of semantic relation indicated by a derivational relation in at least one language (Koeva 2008: 366). Thus, the derivational links in a certain language can be successfully employed for the identification of morphosemantic relations. Furthermore, they can be used for the identification of corresponding semantic relations in other languages, which have different means for lexicalisation. Morphosemantic relations are important for the establishing of the framework of conceptual frames: first, the morphosemantic relations outline subclasses among the word classes: e.g. nouns that can act as human agents, nouns that can act as inanimate agents, etc., and second, the existence of morphosemantic relations between particular verb and noun synsets can serve as a starting point for defining conceptual frames and the enrichment of the semantic network with new semantic relations.

4. Conclusion

Semantic networks organise semantic knowledge and provide effective and integrated access to related data. The structure of the semantic networks allows their enrichment with new semantic relations.

We call conceptual frame a semantic frame whose core frame elements (from the type Entity) are specified for a set of admissible lexical units by means of semantic classes. Our hypothesis is that a semantic class might contain a set of semantic subclasses. For example, within the semantic class [Food] we can introduce the subclass of [Beverage] for verbs like *taste*, *stir*, *sip*, and the subsubclass [Water] for verbs like *drink*, *lap*, etc. Such representation aims to specify the organisation of the concepts into an ontological structure which allows inheritance between the semantic classes down the hierarchy.

The description of conceptual frames encoding the compatibility between subsets of verb and noun classes requires the formulation of their properties and interdependencies. Conceptual frames defining relations between arguments for which semantic restrictions are imposed within the hierarchy of semantic classes provide the theoretical apparatus for the enrichment of semantic networks (as WordNet) with a variety of semantic relations.

Definition of conceptual frames representing the semantic relations between verb synsets from a particular semantic class and noun synsets form particular semantic classes is (largely) language independent and applicable to any wordnet and other semantic networks. An ontological representation is applicable to different languages because semantic classes are (largely) linguistically universal.

5. Organisation of the Book

The book contains three chapters. The chapter *Beyond Lexical and Semantic Resources: Linking WordNet with FrameNet and Enhancing Synsets with Conceptual Frames* (by Svetlozara Leseva and Ivelina Stoyanova) presents the linking of verb inventory in WordNet and FrameNet by mapping the FrameNet semantic frames with the WordNet verb synsets and expanding the mapping to as many synsets as possible by using the relation of inheritance from a hypernym to a hyponym. The chapter *Putting Pieces Together: Predicate-Argument Relations and Selectional Preferences* (by Svetlozara Leseva, Ivelina Stoyanova, Maria Todorova and Hristina Kukova) describes the enhancement of WordNet with semantic relations between verb synsets and noun synsets corresponding to major participants in the predicates' conceptual structure and proposes a typology of selectional preferences (represented as sets of WordNet classes and (sub)trees in the WordNet structure) that verbs impose on the nouns they combine with. The last chapter *Towards Conceptual Frames* (by Svetla Koeva, Tsvetana Dimitrova, Valentina Stefanova and Dimitar Hristov) presents the enrichment of WordNet with verb patterns (predicate – argument structures) from the Pattern Dictionary of English Verbs and the semantic types from the Corpus Pattern Ontology, which semantically describe the elements of the patterns. The PDEV verb patterns were automatically mapped to the WordNet sentence frames, thus adding information about the semantic types of the arguments. The resulting patterns are conceptual frames whose arguments were specified for a set of lexical units – the semantic types assigned to WordNet noun synsets.

References

Baker et al. 1998: Baker, C. F., C. J. Fillmore, J. B. Lowe. The Berkeley FrameNet Project. – In: *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, (COLING-ACL'98), Montreal, Canada, 10–14 August 1998, pp. 86 – 90.

Brachman 1983: Brachman, R. J. What IS-A is and isn't: An analysis of taxonomic links in semantic networks. – In: *Computer*, 16(10), pp. 30 – 36.

Chaffin, Hetmann 1984: Chaffin, R., D. J. Hetmann. The similarity and diversity of semantic relations. – In: *Memory and Cognition*, 13(2), pp. 134 – 141.

Chaflin et al. 1988: Chaflin, R., D. J. Herrmann, M. Winston. An empirical taxonomy of part-whole relations: Effects of part-whole type on relation identification. – In: *Language and Cognitive Processes*, 3(1), pp. 17 – 48.

Cinkova, Hanks 2010: Cinkova, S., P. Hanks. Validation of Corpus Pattern Analysis – Assigning Pattern Numbers to Random Verb Samples, pp. <https://nlp.fi.muni.cz/projects/cpa/CPA_valiman.pdf> [30.05.2020]

Cruse 1986: Cruse, D. A. *Lexical semantics*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge university press.

Di Fabio et al. 2019: Di Fabio, A., S. Conia, R. Navigli. VerbAtlas: a Novel Large-Scale Verbal Semantic Resource and its Application to Semantic Role Labeling. – In: *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, 2019, Association for Computational Linguistics, pp. 627 – 637.

EAGLES 1999: *Preliminary Recommendations on Lexical Semantic Encoding. Final report*. <http://www.ilc.cnr.it/EAGLES96/EAGLESLE.PDF> [30.05.2020]

Fellbaum 1990/1993: Fellbaum, C. English Verbs as a Semantic Net. – In: *International Journal of Lexicography*, Volume 3, Issue 4, Winter 1990, pp. 278 – 301; reprinted in 1993, pp. 40 – 51 . <https://wordnetcode.princeton.edu/5papers.pdf> [30.05.2020]

Fellbaum 1998: Fellbaum, C. (Ed.). *WordNet: An Electronic Lexical Database*. Cambridge (MA): MIT Press.

Fellbaum 2002: Fellbaum, C. On the semantics of troponymy. – In: *The Semantics of Relationships.* Springer, pp. 23 – 34.

Fellbaum et al. 2007: Fellbaum, C., A. Osherson, P. E. Clark. Putting Semantics into WordNet's "Morphosemantic" Links. – In: *Proceedings of the Third Language and Technology Conference*, Poznan, Poland. [Reprinted in: *Responding to Information Society Challenges: New Advances in Human Language Technologies*. Springer Lecture Notes in Informatics], vol. 5603, pp. 350 – 358.

Fillmore 1968: Fillmore, C. J. The case for case. – In: Bach, E., R. Harms (Eds.). *Universals in Linguistic Theory*. New York: Holt, Rinehart, Winston, pp. 1 – 88.

Fillmore 1982: Fillmore, C. J. Frame semantics. – In: *The Linguistic Society of Korea, Linguistics in the Morning Calm*. Seoul: Hanshin, pp. 111 – 137.

Fillmore, Baker 2010: Fillmore, C. J., C. F. Baker. A frames approach to semantic analysis. – In: *Oxford Handbook of Linguistic Analysis.* Oxford: Oxford University Press, pp. 313 – 341.

Hanks 2004: Hanks, P. Corpus pattern analysis. – In: *Proceedings of the 11th EURALEX International Congress,* (EURALEX 2004), Lorient, France, 6–10 July 2004, pp. 87 – 97.

Hanks 2012: Hanks, P. How people use words to make meanings: Semantic types meet valencies. In: A. Boulton and J. Thomas (Eds.) *Input, Process and Product: Developments in Teaching and Language Corpora*. Brno: Masaryk University Press, pp. 54 – 70.

Hanks 2013: Hanks, P. *Lexical Analysis.* Cambridge (MA): MIT Press.
Johnson, Fillmore 2000: Johnson, C., C. J. Fillmore. The FrameNet tagset for frame-semantic and syntactic coding of predicate-argument structure. In: *Proceedings of the Applied Natural Language Processing Conference* (ANLP 2000), pp. 56 – 62.
Kipper et al. 2008: Kipper, K., A. Korhonen, N. Ryant, M. Palmer. A Large-scale Classification of English Verbs. – In: *Language Resources and Evaluation*, 42(1): pp. 21 – 40.
Koeva 2008: Koeva, S. Derivational and Morphosemantic Relations in Bulgarian Wordnet. – In: *Intelligent Information Systems*, XVI, Warsaw, Academic Publishing House, pp. 359 – 389.
Koeva et al. 2004: Koeva, S., S. Mihov, T. Tinchev. Bulgarian Wordnet – Structure and Validation – In: *Journal: Romanian Journal of Information Science and Technology,* volume 7, No. 1–2: pp. 61 – 78.
Koeva et al. 2018: Koeva, S., V. Stefanova, D. Hristov. Semantic Relations in the Scope of Wordnet Multiple Hypernymy. – In: *Foreign Language Teaching,* XLV, 4, 366 – 373.
Landis et al. 1987: Landis, T. Y., D. J. Heimann, R. Chaffin. Development differences in the comprehension of semantic relations. – In: *Zeitschrifl fiir Psychologie*, 195(2), pp. 129 – 139.
Lyons 1977: Lyons, J. *Semantics*. 2 vols. New York: Cambridge University Press.
Miller 1990/1993: Miller, G. Nouns in WordNet: A Lexical Inheritance System. – In: *International Journal of Lexicography*, 3(4), 1990, pp. 245 – 264; reprinted in 1993, pp. 10 – 25 <https://wordnetcode.princeton.edu/5papers.pdf> [30.05.2020]
Miller 1995: Miller, G. A. WordNet: A Lexical Database for English. – In: *Communications of the Association for Computing Machinery*, 38: 11, 39 – 41.
Miller, Fellbaum 2003: Miller, G., C. Fellbaum. Morphosemantic links in Wordnet. – In: *Traitement automatique des langues*, 44.2, pp. 69 – 80.
Miller et al. 1990/1993: Miller, G. A., R. Beckwith, C. Fellbaum, D. Gross, K. J. Miller. Introduction to WordNet: *An On-Line Lexical Database.* – In: *International Journal of Lexicography*, 1990, 3(4), pp. 235 – 244; reprinted in 1993, pp. 1 – 9 <https://wordnetcode.princeton.edu/5papers.pdf> [30.05.2020]
Navigli, Ponzetto 2012: Navigli, R., S. Ponzetto. BabelNet: The Automatic Construction, Evaluation and Application of a Wide-Coverage Multilingual Semantic Network. – In: *Artificial Intelligence*, 193, Elsevier, pp. 217 – 250.
Palmer et al. 2005: Palmer, M., D. Gildea, P. Kingsbury. The Proposition Bank: A Corpus Annotated with Semantic Roles. – In: *Computational Linguistics Journal*, 31:1, pp. 71 – 106.
Piasecki, Koeva 2017: Piasecki M., S. Koeva, S. WordNet relations in the Bulgarian-Polish bilingual perspective. – In: Aleksandrova, T., D. Blagoeva (Eds.). *Proceedings of the International Jubilee Conference of the Institute for Bulgarian Language,* Sofia, 15–16 May 2017, Vol. 1, Sofia: Institute for Bulgarian Language Prof. Lyubomir Andreychin, Bulgarian Academy of Sciences, pp. 285 – 296.
Ruppenhofer et al. 2016: Ruppenhofer, J., M. Ellsworth, M. R. L. Petruck, C. R. Johnson, C. F. Baker, J. Scheffczyk. *FrameNet II*: *Extended Theory and Practice.* <https://framenet2.icsi.berkeley.edu/docs/r1.7/book.pdf> [30.05.2020]
Storey 1993: Storey, Veda. Understanding Semantic Relationships. – In: *The International Journal on Very Large Data Bases,* 2, pp. 455 – 488.
Suchanek et al. 2007: Suchanek, F. M., G. Kasneci, G. Weikum: Yago – A Core of Semantic Knowledge. – In: *Proceedings of the 16th International Conference on World Wide Web*, 16th International Conference on World Wide Web (WWW 2007), Banff, Alberta, Canada, 8–12 May 2007, pp. 697 – 706.

Talmy 1985: Talmy, L. Lexicalization Patterns: Semantic Structure in Lexical Forms. – In: Shopen, T. (Ed.). *Language Typology and Syntactic Description*, vol. 3. Cambridge: Cambridge University Press, pp. 57 – 149.

Urešová et al. 2020: Urešová, Z., E. Fucíková, E. Hajicová, J. Hajic. SynSemClass Linked Lexicon: Mapping Synonymy between Languages. – In: *Proceedings of the Globalex Workshop on Linked Lexicography*, *Language Resources and Evaluation Conference* (LREC 2020), Marseille, 11–16 May 2020, pp. 10 – 19.

Weischedel et al. 2011: Weischedel, R., E. Hovy, M. Marcus, M. Palmer, R. Belvin, S. Pradhan, L. Ramshaw, N. Xue. Ontonotes: A large training corpus for enhanced processing. In: Olive, J., C. Caitlin, J. McCary (Eds.) *Handbook of Natural Language Processing and Machine Translation*. Springer, 54 – 63.

Winston et al. 1987: Winston, M. E., E. Chaffin, D. Herrmann. A taxonomy of part-whole relations. – In: *Cognitive Science*, 11, pp. 417 – 444.

Svetla Koeva
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria
svetla@dcl.bas.bg

BEYOND LEXICAL AND SEMANTIC RESOURCES: LINKING WORDNET WITH FRAMENET AND ENHANCING SYNSETS WITH CONCEPTUAL FRAMES[7]

Svetlozara Leseva, Ivelina Stoyanova

Institute for Bulgarian Language *Prof. Lyubomir Andreychin*
Bulgarian Academy of Sciences

Abstract: This paper sums up the research and findings made towards the linking of two large lexical semantic resources – WordNet and FrameNet (their verb inventory in particular) – by combining the rich conceptual description provided in FrameNet with the vast lexical coverage and structured relational representation in WordNet. The underlying idea is not just to map the two resources to the extent that their overlap allows, but to expand the mapping in such a way as to assign a FrameNet frame to as many synsets as possible, going far beyond the lexical correspondence in the two resources. We suggest that as a first approximation this mapping be implemented by using the relation of inheritance from a hypernym to a hyponym in WordNet to transfer the frame assigned to the parent synset from the parent to the child synset (inheritance-based mapping), and then refine the obtained mapping through a number of procedures.

The paper is divided in two large parts. In the first part we explore the relation of inheritance and other semantic relations as represented in WordNet and FrameNet and how they correspond to each other when the resources are aligned from a theoretical perspective. In addition, we provide a discussion of the implications of these observations with respect to the enhancement of the two resources through the definition of new relations and the detailisation of conceptual frames. The second part of the paper outlines the procedures defined for the validation of the consistency of the inheritance-based mapping of frames and the extension of its coverage which exploit the lexical information in WordNet (synset literals, synset glosses, etc.) and FrameNet (language units and their definitions, the names of the frames, etc.), the structure of the two resources and the systematic relations between synsets in WordNet and between frames in FrameNet. We present a case study on causativity, a relation which provides enhancement complementary to the one using hierarchical relations, by means of linking in a systematic way large parts of the lexicon. We show how consistency checks and denser relations may be implemented on the basis of this relation. In addition, we propose new frames based on causative–inchoative correspondences and in conclusion touch on the possibilities for defining new frames based on the types of specialisation that take place from parent to child synset.

Keywords: *WordNet, FrameNet, conceptual description, verb semantics, linked resources, frame-to-synset mapping*

---

[7] Sections 2 – 5 of this study are a reprint of Sections 2 – 5 of the publication by Leseva and Stoyanova (2019): Leseva, S., I. Stoyanova. Enhancing Conceptual Description through Resource Linking and Exploration of Semantic Relation. – In: *Proceedings of 10th Global WordNet Conference*, 23 – 27 July 2019, Wroclaw, Poland, pp. 229 – 238.

Sections 6 – 8 are a reprint of Sections 4 – 6 of the publication by Stoyanova and Leseva (2019): Stoyanova, I., S. Leseva. A Structural Approach to Enhancing WordNet with Conceptual Frame Semantics. – In: *Proceedings of Recent Advances in Natural Language Processing (RANLP 2019)*, Varna, Bulgaria, Sep 2 – 4, 2019, pp. 629 – 637.

Regarding the description of the resources and their alignment, e.g. the number of synsets with assigned frames, we provide the data reported in the latter (more recent) paper.

1. Introduction

This paper reflects our efforts at linking two large lexical semantic resources, WordNet and FrameNet – in particular their verb inventory – and presents our research findings and results in combining the rich conceptual description provided in FrameNet with the vast lexical coverage and structured relational representation in WordNet.

The underlying idea is not just to link the two resources to the extent that their overlap allows but to expand the mapping in such a way as to assign a FrameNet frame to as many verb synsets as possible. As we aim at full coverage of the 14,103 verb synsets and there are 4,306 unique synset-to-frame mappings already implemented by other research teams which rely primarily on lexical correspondence between units in the two resources, our task means going far beyond this correspondence. We suggest that as a first approximation this mapping be implemented by using the relation of inheritance from a hypernym to a hyponym in WordNet to transfer the frame assigned the parent synset to the child synset (inheritance-based mapping), and then to refine the obtained mapping through a number of procedures.

The paper is divided in two large parts. In the first part (Sections 2 – 5), we start by outlining the principles and procedures of aligning WordNet and FrameNet. The focus is on WordNet as the main lexical-semantic structure (the verbal domain, in particular), which we aim at enhancing with richer linguistic description from FrameNet and VerbNet. We then go on to explore the relation of inheritance and other semantic relations as represented in the two resources and the correspondences between the frame-to-frame relations in FrameNet and the synset-to-synset relations in WordNet. In addition, we provide a detailed discussion of the implications of these observations with respect to the enhancement of the two resources through the definition of new relations and the detailisation of conceptual frames.

The second part of the paper (Sections 6 – 8) outlines the procedures defined for the validation and enhancement of the consistency of the inheritance-based mapping of frames and the extension of its coverage along with a case study on causativity and some proposals for future research. The procedures described in detail exploit both: (i) the lexical information in WordNet (synset literals, synset glosses, etc.) and FrameNet (language units and their definitions, the names of the frames, etc.) and (ii) the structure of the two resources and the systematic relations between synsets in WordNet and between frames in FrameNet. This enhancement is directed to: (a) improving the quality of existing mappings; (b) expanding the mappings' coverage; (c) enhancing the description of frames with additional information obtained from WordNet; (d) proposing structural improvements on the resources based on systemic features (e.g., causativity), including the definition of new conceptual frames; and (e) suggesting further procedures for verification and improvement of the precision.

The aim of the paper is two-fold: (a) from a theoretical perspective, to provide insights into the underlying principles of the representation of semantic information, as reflected in the scope and definition of overlapping or corresponding relations and the relational structure of the two resources, to establish similarities and discrepancies that may come from different semantic construals or from errors; (b) from an applied perspective, to provide directions and techniques for the mutual enhancement and improvement of the two resources, in particular: (i) refining their relational structure; (ii) expanding and refining the accuracy of frame-to-synset mapping using computational procedures based on solid theoretical observations.

After a brief discussion of related work (Section 2), we outline the alignment between WordNet and FrameNet (Section 3) based on existing mappings and procedures for their enhancement and expanding. Section 4 focuses on the theoretical and practical aspects of semantic relations in FrameNet and how they are reflected through respective semantic relations in WordNet, while Section 5 discusses the implications from these observations. Section 6 delves into the procedures for validating, refining and further expanding the mapping between FrameNet and WordNet, whereas Section 7 describes the application of further procedures to the validation and expansion of the frame-to-synset correspondence in the domain of causativity and inchoativity. Section 8 sketches some observations towards frame specialisation. Section 9 concludes the paper by summing up its contributions and highlighting relevant directions of ongoing and future research.

2. Related work

One of the main directions of development of semantic resources is finding ways of uniting their strengths through integrating them and exploiting their features in a complementary way. Mapping of existing semantic resources has been undertaken in a number of works (cf. Section 3.1).

Another line of research in the development and enhancement of the interconnected resources is explicitly linking and generalising existing, but unrelated information in them. A poorly studied direction of research has been the exploration and use of the internal structure of these resources towards their mutual enhancement. One area of research along these lines has been the extension of frame relations by using information from WordNet. Virk et al. (Virk et al. 2016) propose a supervised model for enriching FrameNet's relational structure through predicting new frame-to-frame relations using structural features from the existing FrameNet network, information from the WordNet relations between synsets, and corpus-collected lexical associations. Previously, we have employed features of both relational structures to develop an algorithm for assigning FrameNet frames to WordNet synsets by transferring the relational knowledge for pairs of related synsets to matching lexical units and frames in FrameNet (Leseva et al. 2018).

An interesting theoretical and practical issue arising from the mapping of the 'building blocks' of the two resources is how the underlying relational structures relate and correspond to each other, how they can be mapped to each other, and further explored. Research in this direction is limited (as outlined above) and does not go beyond lexical analysis. Further, evaluation of implemented mappings and the consistency of the data have not been discussed in detail in the literature.

3. Aligning WordNet and FrameNet

Our work relies on two main resources – WordNet and FrameNet, and employs VerbNet as a complementary resource in some tasks related to alignment and verification. We use WordNet (Miller 1995; Fellbaum 1998) as the basic lexical resource. FrameNet (Baker et al. 1998) represents conceptual structures (frames) which describe particular types of objects, situations, etc. along with their participants, or frame elements (Ruppenhofer et al. 2016). Frames are then assigned to lexical units (LUs), e.g. the verb *mature* is assigned the frame Aging with the description 'An Entity is undergoing a change in age typically associated with some deterioration or change in state'. FrameNet is internally structured using a set of relations, which are discussed at length in Section 4. The VerbNet (Kipper-Schuler 2005; Kipper et al. 2008) classes represent formations of verbs with shared semantic and syntactic properties and behaviour organised in a shallow hierarchy.

3.1. Existing mappings

Previous efforts at linking these resources include Shi and Mihalcea (Shi, Mihalcea 2005), Baker and Fellbaum (Baker, Fellbaum 2009), WordFrameNet[8] (Laparra, Rigau 2009; Laparra, Rigau 2010), MapNet[9] (Tonelli, Pighin 2009), and more enhanced proposals, such as the system Semlink[10] (Palmer 2009), which brings together WordNet, FrameNet and VerbNet with PropBank, and its follow-up Semlink+ that brings in mapping to Ontonotes (Palmer et al. 2014). Analysis of the available resources for linking WordNet, FrameNet and VerbNet, as well as procedures for automatically extending the mapping, are presented by Leseva et al. (2018).

These efforts generally suffer from limited coverage and compatibility issues due to multiple release versions of the original resources. Moreover, to the best of our knowledge, no further checks and verification have been performed on the results. This reduces considerably their applicability and further development.

A complementary approach is to exploit the relational structure of the two resources through assigning frames to synsets not only on the basis of direct correspondence between FrameNet LUs and WordNet literals, but also on the basis of the inheritance of conceptual features in hypernym trees and the assignment of frames by inheritance from hypernyms to hyponyms. The main drawback of this approach is that for deeper level WordNet synsets the inherited frames may be underspecified. Our current and prospective work builds upon this paradigm, notably by looking for ways of refining previous proposals (Leseva et al. 2018) through validation which results in enriching the frame structure with systematic relations (e.g. causative, inchoative, etc. frame correspondences). Further, we envisage to define new, more detailed

8 http://adimen.si.ehu.es/web/WordFrameNet [30 May 2020]
9 https://hlt-nlp.fbk.eu/technologies/mapnet [30 May 2020]
10 https://verbs.colorado.edu/semlink/ [30 May 2020]

frames on the basis of more rigid selectional restrictions on frame elements. These procedures are presented in Sections 6 – 8 below.

3.2. Linking procedures

Linking FrameNet to WordNet is not straightforward. There are two principal types of mappings that have already been applied on the lexical resources discussed in Section 3.1: (a) lexical mapping – lexical units (from one resource) have been assigned categories from another, e.g. a FrameNet lexical unit is mapped to a WordNet literal and hence its FrameNet frame is also assigned to the literal (and the synset); and (b) structural mapping – classification categories from one resource have been aligned to categories from another, e.g. a VerbNet class assigned to a synset is linked to a FrameNet frame, so the FrameNet frame is transferred onto the synset. In this way we are able to verify individual mappings by examining the result in terms of the overall structure.

Initially, our mapping is based on three sources of existing lexical mappings: 2,817 direct mappings provided within FrameNet (Baker, Fellbaum 2009), 3,134 from eXtendedWordFrameNet (Laparra, Rigau 2010), and 1,833 from MapNet (Tonelli, Pighin 2009). Structural mapping using VerbNet contributed 1,335 mappings. Overall, there are 4,306 unique WordNet synset to FrameNet frame mappings. The main procedure we apply to improve and extend mapping coverage is based on the relations of inheritance within WordNet. First, we manually verified the frames assigned to 250 out of the 566 root verb synsets: we corrected 75 mappings and assigned valid frames to additional selected 27 root synsets with a large number of hyponyms. We then transferred the hypernym's frame to its hyponyms in the cases where the hyponyms are not directly mapped to FrameNet frames. As a result, we obtained an extended coverage of 13,226 synsets (with an assigned FrameNet frame). With the further defined procedures we aim at improving the quality of this assignment.

The procedures for validation of frame assignments to verb synsets include: (i) manual checks of the assigned frame; (ii) checks for existing but unmapped correspondences between literals and LUs (e.g., by reapplying lexical mapping); (iii) automatic or semi-automatic consistency checks based on correspondences between VerbNet classes (or superclasses) and FrameNet frames; (iv) automatic or semi-automatic consistency checks based on systematic relations within the resources, e.g. causativity (cf. Section 7). If no appropriate frame exists, we propose to posit a new category (and a frame) provided that it is predictable and complies with FrameNet's frame structure. For instance, while Motion is linked to Cause_motion, Self_motion (e.g. {*jump:1, leap:1*} 'move forward by leaps and bounds') does not have a causative counterpart to which verbs such as {*jump:11, leap:4*} 'cause to jump or leap' can be mapped, so we formulate one.

Further, we formulate a detailed proposal for procedures for refining the inheritance-based frame assignment (Sections 6 – 8) based on the theoretical observations of synset-to-synset and frame-to-frame relations and the discussion provided in Sections 4 and 5.

4. Theoretical and practical aspects of semantic relations within FrameNet as reflected in WordNet

FrameNet and WordNet each have its own relational structure which is based on conceptual relations between language units (WordNet) or conceptual representations (FrameNet). The WordNet structure is by far the richer in types and instances of relations; in addition to the conceptual relations it comprises lexical relations, derivational relations and some other relations. Although the relations in the two resources have different number and scope, at least part of them are grounded in similar universal assumptions, which leads to partial overlap, depending on their definition and the specific information in the resources. For instance, there is a clear correspondence between the Inheritance relation in FrameNet and the hypernymy relation in WordNet, to the extent that both represent a modelling of the is-a relation (Ruppenhofer et al. 2016), or between the Causativity relation (FrameNet) and the causes relation (WordNet). Figure 1 presents the process of linking WordNet and FrameNet. In what follows, we are going to explore how the FrameNet frame-to-frame relations translate into WordNet relations (when they do) and to outline the main trends in the correspondence between relations in the two resources.

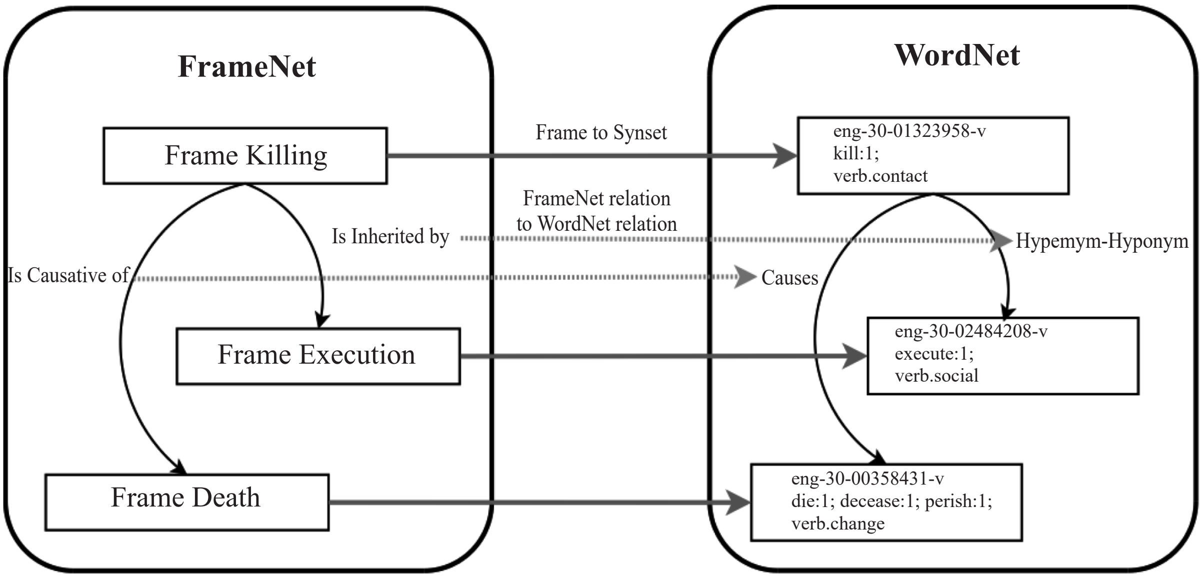

Figure 1: Representation of WordNet to FrameNet linking.

The core part of the data to be examined are pairs or longer chains of WordNet synsets such that: (a) are related through a given WordNet relation, and (b) are assigned FrameNet frames, which are (c) related through a particular FrameNet relation.

The main WordNet relation to be considered is hypernymy, which is the principal tree structure organising relation in the resource. We take into account both direct hypernymy (direct relation between a parent and a child node) and indirect hypernymy (where the hypernym is not a parent of the hyponym but there are intermediate nodes between them). Other relations that emerge from the studied data are: antonymy, also see, causes, verb group, as well as some distant shared hypernyms (i.e. the synsets are in the same tree). Below we present the definition and theoretical grounding of FrameNet relations (Ruppenhofer et al. 2016), along with the observations about their correspondence with WordNet relations.

4.1. Inheritance (Is Inherited by ↔ Inherits from)

Inheritance is defined as the strongest relation in FrameNet; it denotes a relationship between a more general (parent) frame, and a more specific (child) frame in such a way that the child frame elaborates the parent frame. The basic idea, although not always straightforwardly applicable, is that each semantic fact about the parent must correspond to an equally specific or more specific fact about the child (Ruppenhofer et al. 2016, p. 81–82). This means that, generally, there should be a correspondence between entities, frame elements, frame relations and semantic characteristics in the parent and the child frame (Petruck 2015).

Example 1. Frame Killing Is Inherited by frame Execution

Frame: Killing
Core frame elements: Killer; Victim:Sentient; Cause; Means:State_of_affairs; Instrument:Physical_entity
FrameNet definition: A Killer or Cause causes the death of the Victim.
Example synset: {*kill:1*}

Frame: Execution
Core frame elements: Executioner:Sentient; Executed:Sentient
FrameNet definition: An Executioner punishes an individual (Executed) with death as a consequence of some action of the Evaluee (the Reason).
Example synsets: {*execute:1*} (direct hyponym of {*kill:1*}); {*hang:3*} (indirect hyponym)

As per the definition of Inheritance, the configurations of the two frames are similar and the frame elements in the parent frame have correspondences in the child frame, which may be the same or more specific: e.g. Killer has no selectional restrictions, unlike its more specific descendant Executioner (which is specified as Sentient).

Based on this definition, one should expect a considerable overlap between Inheritance and hypernymy: that is, when a pair of WordNet synsets is related through hypernymy and their corresponding frames are related through a frame-to-frame relation in FrameNet, this relation should be Inheritance.

What the data show (Table 1) diverges from this expectation in two ways: (a) there is another frame-to-frame relation which is very strongly favoured for a counterpart of the hypernymy relation, i.e. Using (compare results in Table 1); (b) in a substantial number (20%) of the cases we find out an inverse relationship, i.e. for a hypernym–hyponym pair, the hyponym is assigned the more general (parent) frame, and the hypernym – the child frame in an existing Inheritance relation (the last two

Table 1: WordNet relations hypernymy/hyponymy for different FrameNet relations.

	Is Inherited by		Is Used by		Is Perspec-tivized in		Has Sub-frame(s)		Causative of	
WordNet relation	total	diff.	total	diff.	total	diff.	total	diff.	total	diff.
Direct hypernymy	84	43	67	33	3	2	6	2	13	7
Indirect hypernymy	454	66	576	70	37	2	129	2	41	8
Direct hyponymy	35	22	39	13	0	0	0	0	11	6
Indirect hyponymy	108	21	51	18	0	0	0	0	36	6

rows in Table 1). This is illustrated in Example 2 where the hyponym is assigned the frame Respond_to_a_proposal, while the hypernym receives the child frame Agree_or_refuse_to_act.

Example 2.

Hypernym: {*refuse:1, decline:3*};
Gloss: show unwillingness towards;
Frame: Agree_or_refuse_to_act

Hyponym: {*reject:4, spurn:1*};
Gloss: reject with contempt;
Frame: Respond_to_a_proposal

When looking closely at the data, we find out that in a substantial number of the cases of reversed relation, this is not so much the result of incorrect automatic assignment of frames, as the result of different construal of the conceptual and the lexical domain as the parent and child frames show a high level of similarity. This is the case, though not in all instances, with frame pairs such as Referring_by_name and Labeling, Ingest_substance and Ingestion, Statement and Telling, Statement and Affirm_or_deny, Assistance and Supporting, Change_position_on_a_scale and Proliferating_in_number, among others.

4.2. Using (Is Used by ↔ Uses)

Another hierarchical relation in FrameNet is Using. It is defined as a relationship between two frames where the first one makes reference in a very general kind of way to the structure of a more abstract, schematic frame (Ruppenhofer et al. 2016). The definition has been further specified as a relation between a child frame and a parent frame in which only some of the FEs in the parent have a corresponding entity in the child, and if such exist, they are more specific (Petruck, de Melo 2012); hence, the relation may be viewed as a kind of weak Inheritance (Petruck 2015).

The data confirm that the majority of synsets mapped to FrameNet frames with the Using relation are hypernym-hyponym pairs; also, the numbers for Using are similar to the respective numbers for the Inheritance relation, as shown in Table 1.

Example 3. Frame Placing Is Used by frame Arranging

Frame: Placing
Core frame elements: Agent:Sentient; Cause; Theme:Physical_object; Goal: Goal
FrameNet definition: An Agent places a Theme at a location, the Goal, which is profiled.
Example synset: {*put:1, set:1, place:1, pose:5*}

Frame: Arranging
Core frame elements: Agent:Sentient; Theme:Physical_object; Configuration
FrameNet definition: An Agent puts a complex Theme into a particular Configuration.
Example synset: {*arrange:1, set up:5*}

In Example 3 the child frame and the parent frame to which it refers have similar configurations of elements, with the more specific Configuration (of things) corresponding to Goal (principally a location).

Similarly to Inheritance, cases of inverse assignment of the Using relation, where a hypernym is assigned a child frame, and a hyponym – a parent frame, are also found on a regular basis (12% of the cases) although not as often as with the Inheritance relation. Examples like (4) show that synset members and language units may be mapped to descriptions with different level of specification: in this case {*garage:1*} is construed as more specific in WordNet, but is assigned the more general Placing frame than its hypernym, which receives the frame Storing.

Example 4.

Hypernym: {*store:2*};
Gloss: find a place for and put away for storage;
Frame: Storing

Hyponym: {*garage:1*};
Gloss: keep or store in a garage;
Frame: Placing

The inverse assignment in many of the cases concerns frame pairs which display higher level of similarity and a weaker hierarchical relation. Such frame pairs, though not exclusively, include: Placing – Storing, Abounding_with – Mass_motion, Attempt_suasion – Suasion, Evidence – Explaining_the_facts.

The inverse frame assignment with both Inheritance and Using represents an interesting theoretical issue with respect to the analysis of lexical units (verbs) in terms of their lexical definitions and their conceptual properties.

4.3. Perspective (Is Perspectivized in ↔ Perspective on)

Perspective is defined as similar to, but more specific and restrictive than Using (Ruppenhofer et al. 2016: 82). It indicates that a situation viewed as neutral may be specified by means of perspectivised frames that represent different possible points-of-view on the neutral state-of-affairs.

It follows from this definition that the neutral frame is more abstract than the perspectivised frames and that there should be a great extent of correspondence between the conceptual description and frame elements of the neutral and the perspectivised frames; these features Perspective on shares to a degree with both Inheritance and Using. It is not surprising, then, that this relation may translate as the hypernymy-hyponymy relation (Table 1), and in fact, this is the only WordNet relation that corresponds to it, even though in a very limited way: only 2 pairs of frames are found to be represented by related synsets: Transfer – which is perspectivised in Giving (cf. Example 5) and Hostile_encounter – which is perspectivised in Attack.

Example 5.

Hypernym: {*give:3*};
Gloss: transfer possession of something concrete or abstract to somebody;
Frame: Transfer

Hyponym: {*contribute:2, give:25, chip in:1*};]
Gloss: contribute to some cause;
Frame: Giving

Apart from the actual WordNet relations, we find Perspective on between synsets having a common direct or indirect hypernym, where the same pairs Giving – Transfer and Hostile encounter – Attack are the only two discovered. Only among more structurally distant pairs of synsets do we find other pairs of neutral – perspectivised frames: Transfer – Receiving, Import_export_scenario – Importing, Import_export_scenario – Exporting.

This observation shows that the kind of semantic generalisation underlying the Perspective relation does not correlate well with the WordNet conceptual and lexical relations. In fact, looking more in depth into the data, we find out that synsets related through a WordNet relation may be perspectivised frames of a non-lexical neutral frame. Such example is provided by the antonym pair {*import:1*} (‘bring in from abroad’) – {*export:1*} (‘sell or transfer abroad’): the two synsets are assigned the frames Importing and Exporting, respectively, which perspectivise the neutral Import_export_scenario, and although they have a common hypernym {*trade:1, merchandise:1*}, there is no suitable lexicalisation of the neutral frame. A similar case is presented by other converse (antonym) pairs.

4.4. Subframe (Has Subframe(s) ↔ Subframe of)

Subframe is a relation between a complex frame referring to sequences of states and transitions, each of which can itself be separately described as a frame, and the frames denoting these states or transitions (Ruppenhofer et al. 2016: 83–84). It is also noted that the frame elements of the complex frame may be connected to the frame elements of the subparts, although not all frame elements of one need have any relation to the other. Another feature of this relation is that the ordering and other temporal relationships of the subframes can be specified by the binary Precedence relation.

The definition of Subframe allows for it to correspond to hypernymy, which, apart from 2 instances of also see, is the only WordNet corresponding relation (Table 1), even though it is represented in a very limited way – only 2 pairs of frames are found, Cause_motion – Placing and Cause_motion – Removing (Example 6), and the predominant trend is for non-direct, rather than for direct hypernymy.

Example 6.

Hypernym: {*raise:2, lift:1, elevate:2, get up:3*};
Gloss: raise from a lower to a higher position;
Frame: Cause_motion

Hyponym: {*shoulder:1*};
Gloss: lift onto one's shoulders;
Frame: Placing

In more distant structural relations between WordNet synsets with common non-direct, distant hypernyms, other pairs of frame-to-frame relations are found as well, such as Traversing – Departing, Traversing – Arriving, Intentional_traversing – Quitting_a_place, Self_motion – Quitting_ a_place.

Although Subframe is much better represented through (indirect) hypernymy than Perspective, it shares with it the feature that much like the neutral frame, the complex frame may represent a conceptual structure that does not have a lexicalised correspondence and that it is feasible to look for WordNet relations between subframes of a complex frame (rather than between a complex frame and a subframe). Another supporting example comes from the domain of antonymy – two synsets related by means of the antonymy relation may be assigned subframes of a complex frame, e.g. {*fall asleep:1, dope off:1…*} (Fall_asleep) < antonym > {*wake up:2*} (Waking_up) with respect to Sleep_wake_cycle.

4.5. Precedence (Precedes ↔ Is Preceded by)

This relation holds between component subframes of a single complex frame and provides additional information by specifying the chronological ordering of the states and events (subevents) within a complex event (Ruppenhofer et al. 2016; Petruck 2015). A small number of Precedence instances are found among antonyms (12 pairs) and the majority of the instances are between synsets having a common (direct or indirect) hypernym. The following pairs of frame-to-frame relations are found with antonyms: Placing – Removing, Arriving – Departing, Activity_stop – Activity_ongoing.

Example 7.

Antonym: {*file in:1*};
Gloss: enter by marching in a file;
Frame: Arriving

Antonym: {*file out:1*};
Gloss: march out, in a file;
Frame: Departing

This relation may result in complex structures involving a number of subframes such as the notable example of the Sleep_wake_cycle (Petruck 2015). It does not have a counterpart in the WordNet structure, but it may be transferred, thus bringing an additional dimension of semantic description through linking otherwise unrelated subevents and through specifying their temporal ordering.

4.6. Causation (Causative of) and Inchoativity (Inchoative of)

Causation and Inchoativity are systematic non-inheritance relationships between stative frames and the inchoative and causative frames that refer to them (Ruppenhofer et al. 2016: 85). Obviously, Causation should correspond straightforwardly to the WordNet relation causes. In fact, it does in a small number of cases (30 pairs), which is due to the fact that this relation has not been implemented consistently in FrameNet (Ruppenhofer et al. 2016: 85). It may well be argued that its implementation needs to be enhanced in WordNet as well, as a lot of pairs for which this relation holds have not been linked in the resource. For instance, while the causative and the inchoative sense of *freeze* (see Example 8) are connected through the causes relation, the respective antonym senses have been collapsed in a single synset: {*dissolve:9, thaw:1, unfreeze:1, unthaw:1, dethaw:1, melt:2*} ('become or cause to become soft or liquid').

Example 8.

Synset (causes): {*freeze:4*};
Gloss: cause to freeze;
Frame: Cause_change_of_phase

Synset (is caused by): {*freeze:2*};
Gloss: change to ice;
Frame: Change_of_phase

The lack of the causes relation between causative and inchoative senses is well observed, for instance, in the hypernym trees whose roots are {*change:1, alter:1,* *modify:3*} ('cause to change; make different; cause a transformation') causes > {*change:2*} ('undergo a change; become different in essence; losing one's or its original nature').

There are a considerable number of hypernym-hyponym pairs (see Table 1) that have been assigned the Causation relation. A look at the data shows that these are cases of wrong frame assignment as exemplified in the following case where the causative {*boost:2*} ('give a boost to; be beneficial to') has been assigned the inchoative frame Change_position_on_a_scale instead of the causative frame of the parent synset {*increase:2*} ('make bigger or more'), i.e. Cause_change_of_position_on_a_scale. Such errors in the assignment are commonly found due to the similarity of the formulation of meanings and the common morphological roots of the causative and the inchoative members.

There are 39 correspondences between FrameNet Causative of and WordNet verb group, most of which refer to true causative–inchoative pairs which have not been identified as members of the causes relation in WordNet, as in the following example: {*corrode:1, eat:6, rust:2*} ('cause to deteriorate due to the action of water, air, or an acid'), with the frame Corroding_cause – {*corrode:2, rust:1*} ('become destroyed by water, air, or a corrosive such as an acid'), with the frame Corroding. In these cases, we propose the addition of the more informative causes relation between the respective pairs.

The Inchoativity relation is very poorly represented in the data so we do not consider it herein.

4.7. See also

See also is a relation that has no direct semantic meaning but rather serves to differentiate frames which are similar and confusable (Ruppenhofer et al. 2016: 85, 82). It may be construed in quite different ways, which is reflected in the data, through its mapping to a greater variety of WordNet relations: also see (16 pairs), antonymy (8 pairs), verb group (22 pairs), causes (3 pairs), hypernymy (582 pairs). Example 9 illustrates a See also relation that corresponds to the WordNet also see relation and denotes an unspecified relation of similarity between the Placing and the Filling frame, which represent different profilings of a situation.

Example 9.

Also see: {*put:1, set:1, place:1, pose:5*};
Gloss: put into a certain place or abstract location;
Frame: Placing

Also see: {*put on:7, apply:4*};
Gloss: apply to a surface;
Frame: Filling

The greatest part of the synsets with an actual WordNet relation whose frames are linked by means of See also are related through hypernymy. A typical case is presented in Example 10.

Example 10.

Hypernym: {*search:4*};
Gloss: subject to a search;
Frame: Scrutiny

Hyponym: {*frisk:2*};
Gloss: search as for concealed weapons by running the hands rapidly over the clothing and through the pockets;
Frame: Seeking

The difference between the two frames is stated as one of different primary focus (to the Sought entity or to the Ground)[11]. While this semantic difference is captured by the distinct conceptual structures, it seems to be too fine and does not create a problem in construing {*search:4*} as the hypernym of {*frisk:2*}. Judging from the examples of the hypernym–hyponym pairs and the definition of the frames, the same conclusion is valid for many other pairs of frames, such as Sound_movement – Make_noise, Exchanging – Replacing, Cause_motion – Manipulation, Worry – Experiencer_focused_emotion, Placing – Filling, Motion – Ride_vehicle, among others.

In addition, when examining the See also pairs we find out that many of them are in fact linked through another, more informative relation, e.g. Using: Cause_motion – Bringing, Motion – Operate_vehicle; Inheritance: Motion – Self_motion, Deciding – Choosing; Subframe: Cause_motion – Placing, Cause_motion – Removing.

5. Implications from the observations on semantic relations between frames

The main conclusions that we can make based on the observations so far are:

(1) The internal structure of FrameNet and WordNet is determined primarily by the notion of inheritance (and several non-inheritance relations). In FrameNet this notion is represented by the relations of Inheritance (strong inheritance), Using (weak inheritance) and See also (an unspecified relation of similarity often construable as inheritance), as well as by relations such as Subframe, and Perspective on, although in a limited way. WordNet inheritance is implemented through the hypernymy-hyponymy relation. The comparison between the two structures sheds light on the nature of inheritance and hypernymy, especially in the ways it may diverge from the notion of subsumption. Especially interesting are the cases of inverted relations as they may point to errors in assignment or to a variability in semantic construal.

---

[11] Seeking: A Cognizer_agent attempts to find some Sought_entity by examining some Ground. The success or failure of this activity (the Outcome) may be indicated. NB: This frame should be compared to the Scrutiny frame, in which the primary focus is on the Ground; https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Seeking [30 May 2020]

(2) A practical implication from the comparison refers to the insights into the possible ways of perfecting or enhancing the two resources. We have paid special attention to the way FrameNet relations are translated into WordNet relations. Particularly interesting are cases where relations showing significant similarity in their scope do not correspond in the two resources. Such cases point to peculiarities in the relational structure of the two resources or assignment errors. Inverted relations are also a productive source of information as they point to greater hypernym–hyponym similarity than in straightforward cases and may give clues as to possible collapsing of hierarchical information.

(3) Validation procedures for discovering incorrect assignments of FrameNet frames to WordNet synsets have been proposed on the basis of discrepancies between the two structures through: (i) identifying incompatible relations in the two resources, e.g. FrameNet Causative of and corresponding hypernym-hyponym pairs; (ii) adding relations based on observations, e.g. adding the causes relation between synsets related through verb group; (iii) finding out inaccurately assigned frames by considering pairs of frames not related in FrameNet, but assigned to synsets related through a particular WordNet relation, e.g. Cause_motion – Self_motion, Cause_to_be_dry – Express_publicly, etc.

(4) Suggestion of additional groupings (relations) between synsets on the basis of existing relations. The purpose is to make explicit certain relationships that are not captured (systematically) in WordNet, such as the ones between synsets marked as being subframes of a non-lexicalised complex frame or perspectivised frames of a non-lexicalised neutral frame. The suggestion takes a cue from the way in which temporal relationships between subframes are made explicit through the Precedence relation. For instance, {*fall asleep:1*} and {*wake up:2, awake:1*} are mapped to the Fall_asleep and Waking_up FrameNet frames and are both subframes of the Sleep_wake_cycle. While they are linked through the WordNet antonymy relation, their relationship with synsets representing other subframes of the same scenario remains unaccounted for: {*get up:2, turn out:12*} (Getting_up) and {*sleep:1, kip:1, slumber:1*} (Sleep).

(5) Towards the consistent representation of causativity, we suggest: (a) linking pairs of senses in corresponding causative and inchoative or stative trees, such as the causative and the inchoative change trees (the roots synsets are themselves related through the causes relation); (b) transferring the causes relation to relevant LUs and frames.

(6) The study of the relational structure of the two resources, their overlap and possible improvement has more far-reaching impact with a view to the elaboration of the conceptual structure of verbs undertaken by our team. Based on the properties of the semantic relations in FrameNet and their correlation with hypernymy, we attempt at formulating principles for transferring conceptual information based on the inheritance of features: in particular, configurations of frame elements and imposed selectional restrictions. The observations on Inheritance and Using are especially useful as they shed light on the specialisation that takes place from parent to child: reducing core frame elements by incorporating one of them in the verb meaning – e.g. {*whip:4*} incorporates the Instrument of {*strike:1*}; reducing the scope of the

frame – e.g. {*drive:1*} as a hyponym of {*operate:3*} applies only to land vehicles; profiling a different frame element – e.g. {*rob:1*} profiles the Victim, while its hypernym {*steal:1, rip off:2, rip:4*} profiles the stolen Goods. Among the non-hierarchical relations Causative of and the underrepresented Inchoative of bear importance to the conceptual description as they determine the relations between similar structures with common major frame elements and selectional restrictions. The See also relation denotes similarity between conceptual structures that may very well translate as distinctions between similar configurations of frame elements (as in Example 10) or differences between similar (but not identical) sets of frame elements with similar semantic restrictions.

In the following sections (Section 6 – 8) we propose automatic procedures for the enhancement of the conceptual description of verbs based on the structural properties of WordNet and FrameNet, in particular through refining the inheritance-based mapping.

6. Enhancing the conceptual description of verbs in WordNet

As noted above, whereas FrameNet to WordNet mapping efforts have resulted in the creation of databases of integrated semantic knowledge, they generally suffer from limited coverage as they are restricted to the mapping of the units of the original resources to each other – FrameNet LUs and WordNet synset members (literals), LU definitions and synset glosses, etc. Such a methodology is able to perform mapping in those cases where there is a correspondence between LUs and literals with equivalent or close meaning, but would fail where such correspondence is missing. With 155,287 synonyms in 117,659 synsets and more than 246,577 relations, of which 91,631 are instances of the hypernymy relation as compared with 13,640 LUs and 1,875 frame-to-frame relations in FrameNet, the discrepancy in the size of the data is reflected in the limited coverage of the mappings between synsets and frames. To the best of our knowledge, no further checks and verification have been performed on the mappings, as well.

The approach that we propose in addition to the lexical mapping of units deals with exploring and taking into account the relational structure of the resources (especially the structure of WordNet), particularly the relation of inheritance which ensures the propagation of conceptual and linguistic features down the trees. We employ features of the relational structure in the definition of procedures for the augmentation of the mapping coverage which are aimed at: (i) discovering existing but unmapped relations between synset members and FrameNet frames; and (ii) transferring frames between synsets through relations of inheritance derived from WordNet and FrameNet.

6.1. Enhancing WordNet mappings to FrameNet

As noted above, the proposed approach combines the features used in the direct mapping with the structural properties of WordNet and FrameNet – particularly, the inheritance relations existing between hypernyms and hyponyms in WordNet and the inheritance (and other similar relations) that determine the hierarchical structure of FrameNet. As shown in Section 4 above, although the relations in the two resources have different number and scope, part of them are grounded in similar universal assumptions which leads to partial overlap, depending on their definition and the specificities of the information in the resources.

Apart from the correspondences between FrameNet's Inheritance and other relations and the WordNet hypernymy relation, there are other systematic structural relations which can be applied for the purpose of enhancing the resources. Notable examples are the Causativity relation between frames in FrameNet and the causes relation defined between causative and stative or inchoative verbs in WordNet (cf. Section 4.6, as well as Section 7 below).

6.2. Expanding the mappings based on hierarchical relations in WordNet

Our work relies on the assumption that in a taxonomic structure such as WordNet subordinate nodes inherit the properties of their superordinates, i.e., a hyponym elaborates on the meaning of its hypernym and shares its conceptual and linguistic properties. We propose that if a WordNet synset instantiates a particular FrameNet frame, its hyponyms should (ideally) instantiate the same or a more specific frame which may or may not hold a(n) (inheritance) relation with the more general frame.

This assumption allows us to suggest that in the cases where we are not able to assign a FrameNet frame due to the fact that the coverage of the two resources is non-overlapping and/or other mapping procedures fail, we may resort to assigning the frame of a hypernym to its hyponyms; at worst, the semantic representation will be too general.

Our next concern, therefore, is to define further procedures that would help us to improve and refine the quality of the assignment through inheritance. The sections below present our contributions to enhancing the mappings in the direction of precision and consistency.

6.3. Selection of frames based on the FrameNet and the WordNet structure

We devised two types of procedures aimed at obtaining a more specific mapping: (i) procedures that make use of the conceptual and lexical information and the relational structure in FrameNet; (ii) procedures employing the conceptual and lexical information and the relational structure in WordNet.

As noted above, the first step of assigning a FrameNet frame to a WordNet synset is transferring the frame assigned to the synset's direct or inherited hypernym. The frame so assigned may either appropriately describe the conceptual structure of the literals in the synset, or it may provide a more general description than an optimally informative one. We therefore view this as a default assignment on the basis of which we try to elaborate to the end of discovering a more suitable or specific frame to which to map the synset. When such a frame is found, we validate it manually and assign it to the hyponyms of the synset under discussion, overriding the more general frame as in Example 11.

Example 11.

Synset: {*dress:6, clothe:1, enclothe:1, garment:1*};
Gloss: provide with clothes or put clothes on;
Assigned frame from hypernym: Undergo_change
Suggested frame: Dressing (transferred automatically to 13 out of 15 hyponyms such as {*corset:1*} 'dress with a corset', {*vest:1*} 'dress with a vest', {*overdress:2*} 'dress too warmly')

Below, we describe the procedures proposed and how they make use of the relational structure of FrameNet and WordNet and the following components of the description in the two resources, in particular: (i) WordNet literals (and synsets) and synset-to-synset relations – especially the hypernymy relation, as well as the relations between synsets with a common hypernym (i.e., sister synsets); and (ii) LUs from a particular FrameNet frame (the verbs listed as instantiations of a given frame), the hierarchical frame-to-frame relations: Inheritance, Uses, Subframe, and Perspective, as well as the relation between two frames inheriting from the same frame (i.e., sister frames).

For a synset assigned a frame inherited from its hypernym, we apply the following procedures:

(1) Literal–LU correspondence using FrameNet relations: We check whether any of the synset literals appears as a LU in: (a) the assigned frame (to confirm its validity); (b) more specific frames the frame under discussion is linked to by means of any of the considered frame-to-frame relations (to make it more precise); (c) the sister frames of the assigned frame.

Example 12.

Synset: {*extend:8, expand:4*};
Gloss: expand the influence of;
Assigned frame from hypernym: Cause_change
Suggested frame from (1b): Change_event_duration (LU: *extend*)
Suggested frame from (1c): Cause_expansion (LU: *expand*)

(2) Literal–LU correspondence using WordNet relations: We check whether any of the synset literals appears as a LU in: (a) any of the frames assigned to its hyponyms; (b) any of the frames related to the frames in (a) through frame-to-frame relations; and (c) any of the frames assigned to its sister synsets.

Example 13.

Synset: {*bolster:1, bolster up:1*};
Gloss: support and strengthen;
Assigned frame from hypernym: Cause_change
Suggested frame from (2c): Supporting (LU: *bolster*)

(3) General literal–LU correspondence: We check whether any of the synset literals appears as a LU in any other frame in FrameNet.

Example 14.

Synset: {*exalt:1*};
Gloss: raise in rank, character, or status;
Assigned frame from hypernym: Cause_change
Suggested frame from (3): Judgment (LU: *exalt*)

(4) Keywords: We use keywords (words contained in the FrameNet frame name, plus their derivatives collected from WordNet through the 'eng_derivative' relation), to identify synset literals and/or definitions containing these keywords as candidates to be assigned the frame in question.

Example 15.

Synset: {*clean out:1, clear out:1*};
Gloss: empty completely;
Assigned frame from hypernym: Cause_change
Suggested frame from (4): Emptying (keyword *empty* found in gloss)

(5) Direct similarity: We check the similarity between the gloss of a verb synset and FrameNet LU definitions (even when there is no correspondence between literals and LUs) to identify candidate frames for a given verb synset. We identify: (i) suggested frames related to the one assigned from the hypernym, which are given higher priority; (ii) unrelated suggestions.

Example 16.

Synset: {*gloss over:1, skate over:1, smooth over:1, slur over:1, skimp over:1*};
Gloss: treat hurriedly or avoid dealing with properly;
Assigned frame from hypernym: Intentionally_act
Suggested frame from (5): Avoiding (which Inherits from Intentionally_act); similarity with the definition of LU *skirt.v* 'avoid dealing with'

(6) Indirect similarity: We check the similarity between the glosses of synsets derivationally related to the verb under discussion (as well as the glosses of their hypernyms, which are considered their closest semantic generalisation) and FrameNet LU definitions to identify candidate frames for the verb synset. We identify: (i) suggested frames related to the one assigned from the hypernym, which are given higher priority; (ii) unrelated suggestions.

Example 17.

Synset: {*warn:1*};
Gloss: notify of danger, potential harm, risk;
Assigned frame from hypernym: Telling

Derivationally related synset: {*warning:13*}
Gloss: a message informing of danger;
Suggested frame from (6): Warning (Inherits from Telling); similarity with the gloss of LU *alert.n* 'a message to inform someone of danger; a warning'

Similarity in procedures (5) and (6) is calculated as a cumulative measure based on coinciding terms in the two definitions. Scores of similarity between two words are highest for full match and lower when stemming is applied. Short words (up to length 3) are disregarded and longer words are given more weight. The final score is normalised by the length (in words) of the definitions.

Through these steps 9,341 new suggestions of more specific or other possible frames have been made for 5,661 synsets with automatically transferred hypernym frames – Table 2 shows the distribution of the new suggestions in terms of the types of procedures that have been applied and the distance of the synset from the hypernym whose frame has been inherited.

Table 2: Distribution of frames suggested for synsets with automatic frame assignments from the hypernym (rows (1)-(6) include multiple suggestions for the same synset).

Procedure	# 1-step transfers	# 2-step transfers	# 3-step transfers
(1)	516	231	121
(2)	460	41	17
(3)	1,701	859	145
(4)	1,088	612	27
(5)	1,175	526	202
(6)	1,009	417	194
Unique synsets	3,957	1,388	316

6.4. Discussion on the evaluation

These suggestions need to be manually verified as so far no reliable fully automated verification procedure has been established. Since the main objective is to discover, or suggest, a more precise frame than the one assigned from the hypernym, which is not necessarily wrong but rather may be too general, such evaluation needs to measure the degree of relevance as opposed to precision. Furthermore, it will be highly dependent on the granularity of the frames and their hierarchical organisation. Designing such a measure and its automatisation, if at all achievable, is beyond the scope of this work.

Suggestions, although non-definitive, provide useful pointers to candidate frames and thus are valuable in assisting the manual selection of frames. Only in 203 cases are there multiple suggestions as a result of the procedures, out of which in 177 cases 5 or more different frames are suggested. There are 1,056 synsets for which a suggestion is confirmed at least 2 times from the repeated application of the same or different procedures, out of which 265 cases are confirmed 5 or more times.

Given the task, human judgment is indispensable, especially for frames assigned to synsets higher in the tree as errors propagate down and may result in multiple wrong assignments.

7. Mappings expansion based on causativity and inchoativity as a systemic structural feature

Another direction of expanding the mappings and verifying the information in both FrameNet and WordNet is by employing systematic semantic relations such as causativity. It is a non-hierarchical relation that links stative (e.g. {*lie:2*} 'be lying, be prostrate; be in a horizontal position') or inchoative ({*lie down:1, lie:7*} 'assume a reclining position') verbs with their causative ({*lay:2, put down:2, repose:5*} 'put in a horizontal position') counterparts. The relation provides enhancement complementary to the one using hierarchical relations described above and links in a systematic way large parts of the lexicon.

A considerable part of causative and non-causative pairs are formed with the same root and are thus morphologically similar or identical, e.g. EN *change – change*; RO *schimba – schimba*; BG *promenyam – promenyam se*, which makes them easier to identify. Nevertheless, as noted above, causativity is not consistently encoded in WordNet, and neither is it fully implemented in FrameNet where we have spotted a number of instances of inchoative/stative or causative frames lacking a counterpart in the opposite domain. This means that the verbs instantiating them cannot be appropriately described in FrameNet. Respectively, the mapping of literals instantiating non-defined frames will result in failure of assignment or wrong assignment.

Causativity also has an important application in WordNet and FrameNet data validation and expansion: exploring the assignment of frames from FrameNet to synsets enables us to check the consistency of assignments, by adopting the following logic: (i) in a tree whose root is a causative synset, all the descendants must be assigned a causative frame; (ii) in a tree with an inchoative/stative root all the descendants must be inchoative/stative; (iii) the pairs of causative–non-causative synsets from corresponding trees should be connected to each other through the WordNet causes relation in a consistent way; (iv) the respective pair of causative–non-causative frames assigned to such a pair of synsets should also be related via the Is Causative of relation in FrameNet. The opposite signals either wrong assignment of a frame or inconsistency either in the WordNet data, that is – the encoding of a stative or inchoative verb in a causative tree or vice versa, or in the FrameNet data – missing or wrong relation between frames, undefined frames, etc.

Below we describe the procedures for exploring pairs of causative–non-causative trees and extracting information enabling the validation of assigned frames, as well as the increase of the density of causativity relations within FrameNet and WordNet. Further, we deal with the formulation of new causative or stative and/or inchoative frames.

7.1. Analysis and consistency checks

We have extracted two separate WordNet trees from two root synsets connected by the causes relation (see Table 3): (1) {*change:1, alter:1, modify:3*}, assigned the frame Cause_change; and its corresponding non-causative counterpart (2) {*change:2*}, assigned the frame Undergo_change.

Table 3: Analysed data with respect to causativity (* assignments of the most general frame Cause_change for the causative and Undergo_change for the non-causative).

	Causative (change:1)	**Non-causative (change:2)**
From FrameNet	241	251
Direct hypernym	910	624
Indirect hypernym	577	469
Total	1,728	1,344
General frame*	719	561
In %	41.6%	41.7%

The checks for consistency with regard to (i) – (iv) in Section 7 above, include the following procedures:

(1) Identifying non-causative synsets in the causative tree and causative synsets in the non-causative tree. These mismatches are identified by pattern matching of the gloss or by analysis aimed at establishing whether the manually assigned frame contradicts the position in the tree. 9 such cases have been found in the causative tree (e.g., {*even:6, even out:2*} ‘become even or more even’). Pattern matching in the non-causative tree proved to be unreliable. It identified 120 cases of ‘make’ or ‘cause’ in the gloss, but only a small number of them were causative synsets (e.g., {*break up:3, disperse:1, scatter:1*} ‘cause to separate’). We propose moving each wrongly placed synset (and the subtree rooting from it) to the relevant tree and attaching it to its real hypernym.

Furthermore, there are synsets which combine the causative and the non-causative meaning and thus, create inconsistency in the WordNet structure. We identify such synsets by pattern matching of the gloss since they usually have glosses such as ‘make or become’, ‘cause or become’, ‘cause or undergo’. There are 7 cases in the causative (e.g., {*coarsen:2*} ‘make or become coarse or coarser’) and 5 in the non-causative tree (e.g., {*blacken:1, melanize:1*} ‘make or become black’). We propose that such synsets are split into two and placed at the respective positions in the relevant trees. This is an optimal solution as these concepts are not necessarily expressed by the same lexeme cross-linguistically, and such a split improves the consistency of WordNet.

(2) Identifying non-causative frames assigned to synsets in the causative tree and causative frames assigned to synsets in the non-causative tree. A causative frame is identified based on: keywords such as 'cause' or 'make' in its name or its definition; Agent or Cause / Causer FEs in its conceptual structure; its position as the first member in a Is Causative of relation, etc. A non-causative frame is identified based on: keywords such as 'become' or 'undergo'; lack of Agent or Cause / Causer FEs in its structure; its position as the second member of an Is Causative of relation.

We found 7 non-causative frames in the causative tree (e.g., {*quieten:3, hush:5, quiet:9, quiesce:1, quiet down:1, pipe down:1*} 'become quiet or quieter', Frame: Becoming_silent) and 61 causative frames in the non-causative tree (e.g., {*crush:1*} 'break into small pieces', Frame: Cause_to_fragment). These are clearly either errors in the frame assignment or wrongly encoded synsets as discussed in (1).

(3) Identifying synset pairs connected by the causes relation in WordNet where the causative synset is assigned a non-causative frame or vice versa.

Section 7.2 deals with the enrichment of the two resources with instances of the causative relation.

7.2. Densifying causative relations in WordNet and FrameNet

The causative tree stemming from {*change:1*} and the non-causative one stemming from {*change:2*} were aligned using the WordNet causes relation, resulting in 47 pairs of corresponding synsets – one in each tree. A set of consistency checks showed that there are no crossing relations (i.e., no instances where for a causative hypernym C_1 and its hyponym C_2, and a non-causative hypernym N_1 and its hyponym N_2, C_1 causes N_2 and C_2 causes N_1).

Further procedures were proposed to discover pairs of corresponding causative and non-causative synsets unrelated through the causative relation. These are based on pattern matching of the definition and/or on measuring similarity, as well as on an analysis of the synsets position in the WordNet tree structure, the causative relations in which their sisters, hypernym and hyponyms enter, and the frames assigned to them. On the basis of these linguistic features we have identified 673 possible causative relations between pairs of synsets in the two corresponding trees. After manual validation they may be used to create a more dense structure of causative relations in WordNet, as well as to be extended to frame-to-frame relations in FrameNet.

7.3. Suggesting new frames

New frames are suggested where a suitable causative or non-causative frame is not defined in FrameNet to match its existing counterpart. The missing one is defined using the conceptual description of the available frame. Consider the synset {*age:3*} ‘make older’: we assign it the frame Cause_change and then try to acquire additional classificatory information and, possibly, to find a more specific frame by applying the remaining procedures. We confirm that the synset’s meaning is causative through the keyword procedure (cf. Section 7.2). Another mapping procedure suggests Aging as the corresponding frame. Aging is a non-causative frame denoting the meaning of an entity undergoing a particular kind of change (see Example 18). Since Aging does not have a causative counterpart in FrameNet, we posit such a frame, Cause_to_age. The conceptual structure of stative/inchoative and causative counterparts is distinguished by the presence of a causative subevent in the latter (Van Valin, LaPolla 1997: 109) which is associated with a causative (Agent or an Agent-like) participant (FE). Thus, in the discussed example Cause_to_age is derived from Aging by enriching the set of Aging’s FEs with the frame elements Cause and Agent. In addition, we posit a Causative of relation between Cause_to_age and Aging. In general, causative frames inherit from the abstract frame Transitive_action so we define an Inheritance relation between Transitive_action and Cause_to_age. In such a way the newly-defined relation is integrated into the FrameNet relational structure.

Example 18.

Frame: Cause_to_age
Core frame elements: Agent/Cause; Entity
FrameNet definition: An Agent or Cause causes an Entity to undergo a change in age typically associated with some deterioration or change in state.
Example synset: {*age:3*};
Gloss: make older;

FrameNet relation: Inherits_from
Frame: Transitive_action
Core frame elements: Agent/Cause; Patient
Frame definition: An Agent or Cause affects a Patient.

FrameNet relation: Is_Causative_of
Frame: Aging
Core frame elements: Entity
Frame definition: An Entity is undergoing a change in age typically associated with some deterioration or change in state.
Example synset: {*senesce:1, age:2, get on:7, mature:5, maturate:2*};
Gloss: grow old or older;

The domain of causativity provides an approach at symmetricising large parts of the lexicon both at a horizontal level (same level lexemes in a taxonomic hierarchy) and in depth as the improvements in the higher levels of the lexicon influence the deeper levels as reflected in the procedure of assigning relations by inheritance (Section 6.2).

8. Frame specialisation and relations

The observations on the hierarchical relations, especially on the more populated ones, such as Inheritance, Using and See also, shed light on the specialisation that takes place from parent to child in the taxonomic (inheritance) hierarchy. The changes in the causativity domain deal with including/excluding FEs that correspond to causative subevents in the event structure. The modifications that occur in the conceptual and semantic structure include, but are not limited to the following:

- **Reducing the number of core frame elements by incorporating** one of them in the verb's meaning, e.g. {*whip:4*} incorporates the peripheral FE Instrument ('whip') of {*strike:1*} in the frame Cause_harm assigned to both;
- **Reducing the scope of the frame** through imposing more strict selectional restrictions on the FEs, e.g. {*drive:1*} (Operate_vehicle) as a hyponym of {*operate:3*} (Operating_a_system) applies only to land vehicles while other verbs in the frame impose different restrictions on the FE Vehicle;
- **Profiling a different FE** from the one profiled by the hypernym, e.g. {*rob:1*} (Robbery) profiles the Victim, while its hypernym {*steal:1, rip off:2, rip:4*} (Theft) profiles the stolen Goods;
- **Inclusion/exclusion of a causative/agentive FEs** corresponding to a causative subevent in the respective pairs of frames, e.g. {*break:5*} (Cause_to_fragment) and {*break:2, separate:10, fall apart:4, come apart:1*} (Breaking_apart).

Some of the types of specialisation are currently being studied as a point of departure for defining more narrow-scope frames that would allow for more precise predictions about the selectional restrictions and the syntactic realisation of FEs.

9. Conclusion and future work

This work is an integral part of our research on defining a conceptual framework for encoding semantic relations between verbs (as represented in verb synsets) and relevant semantic classes of nouns (as represented in noun synsets) to the end of creating a relationally densely populated semantic network. This would involve the building of a rich relational structure through defining relations between verbs belonging to particular frames and sets of nouns with particular semantic properties (as reflected in WordNet subtrees, ontological categories, etc.) corresponding to key frame elements in the verb's frame. The linked resource obtained by mapping FrameNet frames and WordNet synsets is a prerequisite for the implementation of this task.

The contribution of the paper consists in a couple of intertwined research avenues. First of all, we propose an implementation of a mapping between WordNet synsets and FrameNet frames by extending existing mappings using the hierarchical structure of WordNet and the concept of inheritance. In addition, considerable improvements on the data have been undertaken including disambiguation of FrameNet frame assignment (selecting a single frame for a given synset, where the mapping has yielded more than one), correction of errors, consistency checks, hypernym assignment to 'orphan' trees, among others.

After obtaining the initial mapping we go on to undertake a theoretical study of FrameNet frame relations and their correspondences in WordNet in terms of their theoretical grounding, their definition and scope. This has led us to discover some underlying principles of the structuring of the two resources, along with existing but non-explicit relations in one of them that are mappable to the other. This has made an impact on the formulation of procedures for a more reliable frame assignment using semantic inheritance.

This theoretical research has informed the devising of a semi-automatic methodology (automatic assignment complemented by manual verification at several stages) aimed at a structured and consistent enrichment of the two resources, which presents an important step towards its automatisation.

Part of our future work is directed to the definition of new and specialised frames where the existing ones prove insufficiently detailed or appropriate to describe the conceptual semantics of verbs and verb classes.

Another line of research that we intend to undertake is the enrichment of FrameNet by extending its lexical coverage on the basis of the expanded mapping to synsets. Verbs which have no correspondence among LUs (or no correspondence in a given frame) but belong to synsets that have been successfully mapped to FrameNet frames, will be suggested as possible LUs to be included in the respective frame(s).

A major vein of promising ongoing research is to define appropriate selectional restrictions on FEs and to implement them as semantic relations between a verb synset and a set of noun synsets that satisfy these restrictions. In such a way we will enrich WordNet with relations between verbs and nouns corresponding to participants in their conceptual structure, particularly ones realised as arguments and adjuncts.

The obtained linked and enriched resource may have a considerable impact on the development of methods for identification of predicate-argument structure in text, which in turn will facilitate the development of new methods for frame verification and consistency checks on FrameNet and WordNet. To this end, it will be made available to the research community. As the description of verbs' conceptual structure is largely language independent, the enriched description is applicable cross-linguistically.

References

Baker et al. 1998: Baker, C. F., C. J. Fillmore, J. B. Lowe. The Berkeley FrameNet Project. – In: *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (COLING-ACL'98)*, 10 – 14 August 1998, Montreal, Canada, pp. 86 – 90.

Baker, Fellbaum 2009: Baker, C. F., C. Fellbaum. WordNet and FrameNet as Complementary Resources for Annotation. – In: *Proceedings of the Third Linguistic Annotation Workshop (ACL-IJCNLP'09)*, ACL, Stroudsburg, PA, USA, pp. 125 – 129.

Fellbaum 1998: Fellbaum, C. (Ed.) *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

Kipper et al. 2008: Kipper, K., A. Korhonen, N. Ryant, M. Palmer. A Large-scale Classification of English Verbs. – In: *Language Resources and Evaluation*, vol. 42 (1), pp. 21 – 40.

Kipper-Schuler 2005: Kipper-Schuler, K. *VerbNet: A Broad-coverage, Comprehensive Verb Lexicon*. PhD. Thesis. Computer and Information Science Dept., University of Pennsylvania. Philadelphia.

Laparra, Rigau 2009: Laparra, E., G. Rigau. Integrating WordNet and FrameNet using a Knowledge-based Word Sense Disambiguation Algorithm. – In: *Proceedings of Recent Advances in Natural Language Processing* (RANLP'09), 14 – 16 September 2009, Borovets, Bulgaria, pp. 208 – 213.

Laparra, Rigau 2010: Laparra, E., G. Rigau. eXtended WordFrameNet. – In: *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010)*, 17 – 3 May 2010, Valletta, Malta, pp. 1214 – 1219.

Leseva et al. 2018: Leseva, S., I. Stoyanova, M. Todorova. Classifying Verbs in WordNet by Harnessing Semantic Resources. – In: *Proceedings of the Third International Conference on Computational Linguistics in Bulgaria (CLIB 2018)*, 28 – 29 May 2018, Sofia, Bulgaria, pp. 115 – 125.

Miller 1995: Miller, G. A. WordNet: A Lexical Database for English. – In: *Communications of the ACM*, 38 (11), pp. 39 – 41.

Palmer 2009: Palmer, M. Semlink: Linking PropBank, VerbNet and FrameNet. – In: *Proceedings of the Generative Lexicon Conference*, pp. 9 – 15.

Palmer et al. 2014: Palmer, M., C. Bonial, D. McCarthy. SemLink+: FrameNet, VerbNet and Event Ontologies. – In: *Proceedings of Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929–2014)*, Baltimore, USA, ACL, pp. 13 – 17.

Petruck 2015: Petruck, M. R. The Components of FrameNet. *Tutorial at NAACL 2015.* International Computer Science Institute, Berkeley, CA. <http://naacl.org/naacl-hlt-2015/tutorial-framenet-data/FNComponentsMRLP.pdf> [01.06.2020].

Petruck, de Melo 2012: Petruck, M. R., G. de Melo. Precedes: A Semantic Relation in FrameNet. – In: *Proceedings of the Workshop on Language Resources for Public Security Applications*, pp. 45 – 49.

Ruppenhofer et al. 2016: Ruppenhofer, J., M. Ellsworth, M. R. L. Petruck, C. R. Johnson, C. F. Baker, J. Scheffczyk. *FrameNet II: Extended Theory and Practice* (Revised November 1, 2016). <https://framenet2.icsi.berkeley.edu/docs/r1.7/book.pdf> [01.06.2020]

Shi, Mihalcea 2005: Shi, L., R. Mihalcea. Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robust Semantic Parsing. – In: Gelbukh, A. (Ed.). *Computational Linguistics and Intelligent Text Processing (CICLing 2005)*. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, vol. 3406, pp. 100 – 111.

Tonelli, Pighin 2009: Tonelli, S., D. Pighin. New Features for FrameNet – WordNet Mapping. – In: *Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL'09)*, Boulder, USA.

Van Valin, LaPolla 1997: Van Valin Jr., R. D., R. J. LaPolla. *Syntax: Structure, Meaning and Function*. Cambridge University Press.

Virk et al. 2016: Virk, S. M., P. Muller, J. Conrath. A Supervised Approach for Enriching the Relational Structure of Frame Semantics in FrameNet. – In: *Proceedings of the 26th International Conference on Computational Linguistics (COLING 2016)*: *Technical Papers*, Osaka, Japan, pp. 3542 – 3552.

Svetlozara Leseva
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria
zarka@dcl.bas.bg

Ivelina Stoyanova
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria
iva@dcl.bas.bg

PUTTING PIECES TOGETHER: PREDICATE-ARGUMENT RELATIONS AND SELECTIONAL PREFERENCES

Svetlozara Leseva, Ivelina Stoyanova, Maria Todorova, Hristina Kukova

Institute for Bulgarian Language *Prof. Lyubomir Andreychin*
Bulgarian Academy of Sciences

Abstract: This paper presents the work on enhancing WordNet with semantic relations between verb synsets and classes of noun synsets corresponding to major participants in the predicates' conceptual structure. We first provide the theoretical background and motivation for the study and discuss the integration of the three complementary semantic resources – WordNet, VerbNet and FrameNet – which we then use to the end of devising a framework for enriching the relational structure of WordNet with a system of predicate-argument and predicate-adjunct relations. We pay particular attention to the analysis of the relations of inheritance between conceptual frames in FrameNet and between frame elements (the elements of these conceptual frames) which results in the elaboration of a hierarchy that is then translated as a set of relations and relation subtypes between predicates and the main elements in their conceptual structure. The conceptual frames with their corresponding frame elements and selectional restrictions are assigned to verb synsets in WordNet.

We then go on to propose a typology of selectional preferences that verbs impose on the nouns they combine with. Using restrictions that have already been defined in FrameNet and VerbNet as well as other semantic information from the three resources, we propose a unified and extended set of selectional preferences represented as sets of WordNet classes and (sub)trees in the WordNet structure. The model is illustrated by a case study of the relation *Theme* and its subtypes.

Key words: c*onceptual structure, predicate-argument relations, selectional preferences, Frame semantics, WordNet, VerbNet, FrameNet*

1. Introduction

The research presented in this paper aims at establishing the theoretical motivation behind the methods we have adopted towards the enhancement of WordNet with information about the conceptual structure of verbs. After a brief overview of the three resources, WordNet, FrameNet and VerbNet, as well as the mappings between them (Section 2), we focus on the main theoretical principles of our research (Section 3). The enrichment of WordNet with semantic relations which capture the selectional restrictions between a verb synset and classes of noun synsets denoting the verb's arguments and possibly some adjuncts are laid out in Sections 4 – 6. An additional goal of this study is to define a framework for formalising these restrictions on the basis of a detailed conceptual representation of verbs (adopted from FrameNet) and the description of prominent semantic features that predetermine the selectional preferences of these verbs. To this end, we harness information from the three above-mentioned resources which informs the typology of selectional preferences proposed in Section 7. We illustrate this framework in a case study of the relation *Theme* which encompasses a range of specific relation types and varieties (Section 8).

2. Used resources

As noted above, the proposed analysis relies on the use of the combined information available in WordNet, FrameNet and VerbNet, which results in a rich representation of paradigmatic and syntagmatic aspects of lexical semantics (Baker, Fellbaum 2009).

2.1. WordNet

WordNet[12] (Miller 1995; Fellbaum 1998) is a large lexical database that represents comprehensively conceptual and lexical knowledge in the form of a network whose nodes denote cognitive synonyms (synsets) interconnected through a number of conceptual-semantic and lexical (including derivational) relations such as hypernymy, meronymy, etc. The main relation that determines WordNet's structure is the relation of hypernymy. It is especially useful in our work as many of the restrictions are defined as treelike structures with a common hypernym (subtrees).

Of the three resources, WordNet provides the most coarsely-grained semantic division in terms of a set of language-independent semantic primitives (semantic classes) assigned to all the nouns and verbs in the resource. The nouns are categorised into 25 groups, such as noun.act (acts or actions), noun.artifact (man-made objects), noun.person (persons). The verbs fall into 15 groups, such as verb.change (verbs describing change in terms of size, temperature, intensity, etc.) and verb.cognition (verbs of mental activities or processes)[13].

The original Princeton WordNet has given rise to the construction of similar networks, including wordnets for Bulgarian and for other Balkan and Slavic languages, among others. The corresponding synsets in many of the individual wordnets are related to each other through unique interlingual identifiers. In such a way, the lexical and conceptual knowledge is aligned cross-linguistically, which makes it possible for inter-lingual studies of semantic and syntactic correspondences to be conducted. The observations presented below are transferable to the Bulgarian WordNet (BulNet) through its alignment with Princeton WordNet.

2.2. FrameNet

FrameNet[14] (Baker et al. 1998; Baker 2008) is a resource which couches lexical and conceptual knowledge in the apparatus of frame semantics. Frames are conceptual structures describing particular types of objects, situations, or events along with their components, called frame elements, or FEs (Baker et al. 1998; Ruppenhofer et al. 2016). Depending on their status, FEs may be core, peripheral or extra-thematic

[12] https://wordnet.princeton.edu/; Princeton WordNet may be explored online at: http://wordnetweb.princeton.edu/perl/webwn.

[13] The division of the nouns and verbs into WordNet lexicographic files (reflecting the semantic primitive distinction) along with short definitions of the primitives are available at: https://wordnet.princeton.edu/documentation/lexnames5wn.

[14] FrameNet frames and relations may be accessed at: https://framenet.icsi.berkeley.edu/fndrupal/.

(Ruppenhofer et al. 2016). For our purposes, we deal particularly with core FEs, which instantiate conceptually necessary components of a frame, and which in their particular configuration make a frame unique and different from other frames. Semantic restrictions in FrameNet are more fine-grained than those in WordNet and are represented as semantic types that are assigned to relevant frame elements. FrameNet's theoretical framework has been successfully adopted for various languages (see Baker 2008 for a discussion), thus corroborating the assumption of the universality of the conceptual description proposed in the resource. The framework has been adopted for Bulgarian and extended into an even richer model which accounts for language-specific features, including verb aspect, semantic and syntactic diatheses and syntactic alternations, among others (Koeva 2010).

FrameNet frames are related in a hierarchical network by means of a number of frame-to-frame relations. The semantics and scope of these relations with a view to their WordNet counterparts are studied in-depth in Leseva and Stoyanova (Leseva, Stoyanova 2020, this volume). In this paper, we use primarily the inheritance between pairs of frames and corresponding frame elements in them to the end of exploiting the so-formed hierarchy of frame elements and semantic features (Section 6 below).

2.3. VerbNet

VerbNet (Kipper-Schuler 2005; Kipper et al. 2008) is a hierarchical network of English verbs which represents their syntactic and semantic patterns[15]. It is organised into 274 verb classes extending Levin's classification (Levin 1993) through refining and adding subclasses so as to provide better syntactic and semantic coherence among members of a class. VerbNet explicitly projects semantic relations onto syntactic structures and encodes information about thematic roles, arguments' selectional restrictions and syntactic frames. While the syntactic dimension of the resource is more specific to English, the selectional restrictions provide well-motivated semantic generalisations which as a whole complement the FrameNet semantic types. Where relevant, we have combined them with the semantic information from the other two resources. These selectional restrictions are described in Section 7 below.

2.4. Mapping the resources

A successful combination of the three resources should be able to exploit and enhance their individual strengths: the extensive lexical coverage and the branched and rich relational structure of WordNet, the detailed conceptual description of the combinatorial potential of lexical units supplied by FrameNet and the selectional restrictions and semantic and syntactic generalisations encoded in VerbNet.

The three resources have been mapped automatically using existing mappings or newly designed procedures in such a way that WordNet synsets are assigned corresponding verb classes from VerbNet and frames from FrameNet. The alignment

[15] https://verbs.colorado.edu/verbnet/ [30 May 2020]

was based on previously implemented mappings which have been supplemented and partially validated[16]. As in reality the limited overlap between the resources impacts the results and usability of their alignment, efforts have been particularly directed towards the expansion of the mapping between WordNet and FrameNet (which we focus on currently) with a view to obtaining maximum coverage and overcoming the sparsity of the overlap. For an overview of the work done in this vein of research cf. Leseva and Stoyanova (Leseva, Stoyanova 2019; Leseva, Stoyanova 2020, this volume). The mapping between synsets and frames proposed there expands on previous ones; the data obtained from this alignment were used in the analysis and the definition of relations and selectional preferences below.

3. A theoretical overview and related work

The aim of this research is the formulation of a detailed inventory of semantic relations corresponding to elements of the conceptual structure of predicates and reflecting the semantics and the selectional restrictions imposed on their arguments (as well as some non-argument participants), which we propose to be encoded between verbs and the appropriate classes of nouns in WordNet.

The definition of these relations facilitates the consistent integration of information about the conceptual structure of verbs into the WordNet[17] relational structure in such a way that each predicate–argument or predicate–adjunct relation is couched in terms of a semantic relation between a verb belonging to a synset and a set of noun synonym sets that correspond to the verb's arguments (or adjuncts); the noun synsets are usually organised in a WordNet subtree structure reflecting the valid semantic preferences imposed by the verb. The formulation of the relations is based on the analyses of relevant information from the three employed semantic resources – the semantic verb classes in WordNet, the semantic roles and selectional restrictions in VerbNet and the conceptual frames and their elements and semantic types in FrameNet.

The semantic relations between predicates and their arguments have been studied within the frameworks of different theoretical approaches based mainly on: (a) syntactic features and behaviour (Levin 1993; Pinker 1989; Goldberg 1994, among others), (b) thematic structure (Chafe 1970; Longacre 1976; Foley, Van Valin 1984; Van Valin, LaPolla 1997), (c) frame semantics (Fillmore 1982). These theoretical accounts have been applied explicitly or implicitly in the construction of the three semantic resources involved in our study.

[16] We use the mapping of the VerbNet 3.3 verb classes and the synsets in WordNet (http://verbs.colorado.edu/verbnet/index.html) as well as two types of mappings of the frames in FrameNet and the synsets in WordNet: indirectly via SemLink (https://verbs.colorado.edu/semlink/) and directly through the system described in Laparra and Rigau (Laparra, Rigau 2010; http://adimen.si.ehu.es/web/WordFrameNet).

[17] For this research we use data from the Princeton WordNet (https://wordnet.princeton.edu/) and the Bulgarian WordNet (http://dcl.bas.bg/bulnet/) [30 May 2020]

The understanding that the mapping of semantic resources leads to their mutual enrichment and increases their value in various applications underlies the work of many researchers. Notable efforts in this field include the work of Shi and Mihalcea (Shi, Mihalcea 2005) on the mapping of the three resources under discussion, the implementation of FrameNet-to-WordNet mappings, such as WordFrameNet[18] by Laparra and Rigau (Laparra, Rigau 2010), MapNet[19] by Tonelli and Pighin (Tonelli, Pighin 2009), the alignments proposed by Fellbaum and Baker (Fellbaum, Baker 2008) and Ferrandez et al. (Ferrandez et al. 2010), among others. More recent and enhanced proposals have been made – in particular, the system Semlink (Palmer 2009), which brings together WordNet, FrameNet and VerbNet with PropBank, and its follow-up Semlink+ that, in addition, includes mapping to Ontonotes (Palmer et al. 2014).

Even more recently, the SynSemClass lexicon[20] has marked a distinguishable effort towards combining a rich semantic description of verbs with external semantic resources in order to create a multilingual contextually-based verb lexicon. As described by Urešová et al. (Urešová et al. 2020), the aim of the lexicon is to provide a resource of classes of verbs that compares their semantic roles as well as their syntactic properties. In addition, each entry is linked to FrameNet, WordNet, VerbNet, OntoNotes and PropBank, as well as the Czech VALLEX to the end of building a type ontology of events, processes and states.

VerbAtlas[21], proposed by Di Fabio et al. (2019), is a hand-crafted lexical semantic resource which represents synsets as clusters with prototypical argument structures presented as frames, to a large extent inspired by VerbNet roles and semantic restrictions. In a similar spirit as the one proposed in this paper, the restrictions in VerbAtlas are defined as selectional preferences mapped to classes of noun synsets in WordNet.

Another notable study is the enrichment of the Princeton WordNet with information about the verb patterns (predicate-argument structures) adopted from the Pattern Dictionary of English Verbs (PDEV) and the semantic types proposed in the Corpus Pattern Ontology (CPA) which describe the arguments of these patterns. The PDEV verb patterns are automatically mapped to the WordNet sentence frames; as a result, the authors obtain conceptual frames of verbs whose arguments are specified for a set of lexical units represented as WordNet noun synsets (Koeva et al. 2020, this volume).

The definition of predicate–argument (and predicate–adjunct) relations within the WordNet framework draws on the already encoded morphosemantic relations between derivationally related verbs and nouns in WordNet (cf. Koeva 2008; Koeva et al. 2016 for Bulgarian), most of which describe relations between predicates and synsets denoting participants in their conceptual structure and correspond to well-established semantic roles.

[18] https://adimen.si.ehu.es/web/WordFrameNet [30 May 2020]
[19] https://hlt-nlp.fbk.eu/technologies/mapnet [30 May 2020]
[20] https://ufal.mff.cuni.cz/synsemclass [30 May 2020]
[21] http://verbatlas.org/ [30 May 2020]

The work described in this paper extends these morphosemantic relations both in number (introducing new relations or subtypes of existing ones) and in scope by defining them independently of the morphosemantics and founding them on entirely semantic grounds. We take as a point of departure the inventory of verb senses in WordNet, which we have further enriched with conceptual information by aligning them with conceptual frames in FrameNet through the WordNet-to-FrameNet mapping, the latter resource being the one providing the most detailed and language-independent framework for conceptual description.

We then use the semantic information in WordNet, FrameNet and VerbNet to define relations between the predicates and the main participants in their conceptual frame, on the one hand, and selectional preferences imposed on the nouns entering in these relations, on the other, casting these preferences as WordNet classes as proposed in previous research (Agirre, Martinez 2002; Koeva 2010). In addition to enriching WordNet with conceptual information from FrameNet and with semantic relations to noun classes, we aim at achieving consistency of representation, where systemic relations such as Inheritance (a relation between frames and between frame elements in FrameNet, as well as between synsets in WordNet) are employed in order to validate the assigned relations and their hierarchy.

4. A methodology for enriching the WordNet relational structure

The enrichment of WordNet's relational structure may be viewed as carrying out the following tasks: (i) selecting and defining a set of semantic relations on the basis of the analysed data; (ii) defining the selectional preferences for each relation with respect to the verb classes and possibly individual verbs (verb synsets) that enter into this relation, and consequently – defining the set of noun synsets (as a WordNet subtree, where possible) that correspond to these restrictions.

The resources used in the study encode semantic information of different levels of granularity; respectively the corresponding elements of the description in each of them represent different levels of semantic generalisation. WordNet provides an ontological characterisation of verbs and nouns according to their semantic primitive (see Section 2.1). VerbNet uses an inventory of well-established semantic roles, such as *Agent, Experiencer, Theme, Patient, Goal, Source, Result, Product, Stimulus*, etc. The definition of the roles, along with a shallow hierarchy describing the inheritance of semantic features between roles are laid out in the special Guidelines[22]. The semantic roles are further specified by means of a set of selectional restrictions that denote the existence (+) or absence (–) of properties, such as [SUBSTANCE], [ANIMATE], [MACHINE], etc., and may be combined with operators, such as | (OR) and & (AND). For instance, a common restriction across verb classes states that an *Agent* may be an animate being or an organisation: *Agent* [+ANIMATE | +ORGANIZATION].

[22] https://verbs.colorado.edu/verb-index/VerbNet_Guidelines.pdf [30 May 2020]

We use this information in two ways. First of all, we implemented correspondences between VerbNet semantic roles and FrameNet frame elements. This comparison helped us cross-check the definitions and scope of the roles and the frame elements to the end of understanding the semantic grounds for each of them and making an informed choice with respect to the description of predicate–argument and predicate–adjunct relations. In addition, we compared the selectional restrictions in VerbNet with those defined by the semantic primitives in WordNet and the semantic types of the frame elements in FrameNet in light of selecting features that provide robust generalisations about the semantic classes of nouns that meet the selectional preferences imposed by a verb on its arguments; in such a way we tried to ensure that the preferences are relevant across languages and resources.

Part of the frame elements correspond straightforwardly to recognised semantic roles. However, as FrameNet adopts a more fine-grained approach to the description of the conceptual structure of lexical units, many frame elements are defined more narrowly and represent specialisations of well-known semantic roles. We derive these specialisations by employing the inheritance relation between frames and between frame elements. As a result, we build a hierarchy of semantic relations between predicates and frame elements in their conceptual structure which is essential for the definition of selectional preferences and restrictions.

Consider for instance the verbs which evoke (and whose conceptual structure is described by) the frame Text_creation: *author*, *chronicle*, *compose*, *draft*, *jot*, *print*, *type*, *write*, among others. The frame elements corresponding to the person who creates the text and the textual matter created are defined as *Author* and *Text*, respectively. On the basis of the relation of inheritance defined between the more abstract frame Intentionally_create and the more specific one Text_creation (the main inheritance relations between frames are Inherits_from and its inverse relation Is_inherited_by, as well as Uses and its inverse Is_Used_by, also known as weak inheritance, cf. Leseva, Stoyanova 2020, this volume), we may infer that *Author* is a specialisation of the FE *Creator*, whereas *Text* is a specialisation of the FE *Created_entity* in the frame Intentionally_create. Using the relation of inheritance defined between Intentionally_create (relevant core FEs[23]: *Creator, Created_entity*), on the one hand, and more specific frames in addition to Text_creation, such as Building (relevant FEs: *Agent, Created_entity*), Cooking_creation (relevant FEs: *Cook, Produced_food*), Create_physical_representation (relevant FEs: *Creator, Representation*), Manufacturing (relevant FEs: *Producer, Product*), on the other, we derive the relation of inheritance between the corresponding FEs: *Creator* > *Author* | *Cook* | *Producer*; *Created_entity* > *Produced_food* | *Product* | *Representation* | *Text*. Going further and analysing the frames from which Intentionally_create inherits its properties, as well as the correspondences with the respective verbs in VerbNet, we may derive the following generalisations

[23] "Relevant" in this case means that we consider only those (core) FEs of a frame that are related through inheritance to FEs in the more general frame. In some cases the more specific frame might have additional (core) frame elements which are not relevant for the study of frame element specialisation.

about the hierarchical relations between more abstract and more specialised roles/FEs: *Agent* > *Creator* > *Author* | *Cook* | *Producer*; and *Patient* > *Created_entity* > *Text* | *Produced_food* | *Product* | *Representation*.

The detailed exploration of pairs of frames and frame elements related through inheritance in FrameNet as well as of pairs of FrameNet frame elements and semantic roles in VerbNet has made it possible to obtain information about: (i) the correspondences between semantic roles and frame elements; and (ii) possible specialisations of well-known semantic roles based on more fine-grained semantic distinctions underlying the formulation of conceptual frames in FrameNet.

We then go on to define predicate–argument (and predicate–adjunct) relations based on the generalisations discussed above. Basically, if there is a given FE in a frame's conceptual description, the verbs evoking this frame stand in the same relation with nouns satisfying particular semantic preferences, e.g., an *Agent* FE in a conceptual frame is related through the relation *Agent* with the verbs of this frame. As we derived a hierarchy between more general and more specific FEs, we were able to create a typology of each general relation and group the more specific ones into categories. For instance, we can lump together *Author*, *Producer* and *Cook* with their superordinate *Creator*, which affords a greater level of abstraction, while preserving information about the prototypical *Agent* of verbs of creation. We can opt for an even more abstract level of description or preserve the original level of granularity. The shallow hierarchy allows us to revise the generalisations or to make new ones without loss of information; it is also used when defining the selectional preferences for the classes of noun synsets that correspond to the arguments or adjuncts of verbs. For instance, *Cooks* and *Authors* would typically be persons, while *Producers* may be persons or organisations (companies).

We decide on the definition of the predicate–argument and predicate–adjunct relations on the basis of analysis and unification of the data in compliance with the following principles:

(i) Level of abstraction of the relation measured in terms of the number of frames to which the respective frame element (candidate relation) is a member and the number of verb semantic primitives across which it is found. Elements participating in a small number of frames generally do not qualify for a separate relation.

(ii) Correspondence with verbs' semantic primitives. A relation which corresponds to a semantic primitive is considered to be the prototype of the predicate–argument relation in which the verbs from the respective class enter, for instance *Communicator* is the prototypical agentive relation for verbs of communication, *Creator* – the prototypical agentive relation for verbs of creation.

(iii) Semantic distinctness of the relation in terms of its ontological membership, its definition and selectional preferences.

The analysis in this study is based on 4,522 verb synsets in WordNet to which frames have been assigned and consequently manually validated. The mapping involves 465 different frames with a total of 405 unique frame elements, a large part of which have been individually validated too. A system of 93 unique general restrictions on the selectional preferences imposed on frame elements has been applied to define the noun classes that the analysed verbs enter into relations with (Section 7).

5. Modelling semantic relations based on the verb argument structure

Predicate–argument and predicate–adjunct relations have already been introduced in WordNet in the form of morphosemantic relations, such as *Agent, Instrument, Result, Undergoer*, etc. They are defined between derivationally related English verb–noun pairs in the Princeton WordNet (Fellbaum et al. 2009) and have been automatically transferred to the equivalent synsets in BulNet regardless of whether the Bulgarian counterparts are derivationally related (Koeva et al. 2016), e.g. EN: *{cook:3} – {cook:1}*, BG: *{gotvya:2, sgotvyam:2, sgotvya:2, prigotvyam:2, prigotvya:2} – {gotvach:1}*, Relation: *Agent*; EN: *{heat:1, heat up:2} – {heater:1, warmer:1},* BG: *{nagoreshtyavam:1, nagoreshtya:1, nazhezhavam:1, nazhezha:1, nagryavam:2, nagreya:3} – {nagrevatel:1, nagrevatelen ured:1}*, Relation: *Instrument.* Not all morphosemantic relations denote such argument or adjunct relations, consider EN: *{cook:3} – {cooking:1}*, BG: *{gotvya:2, sgotvyam:2, sgotvya:2, prigotvyam:2, prigotvya:2} – {gotvene:1, gotvarstvo:1, sgotvyane:1, prigotvyane:3}*, which denotes an eventive meaning.

Morphosemantic relations as defined in WordNet link a verb with a noun which is a prototypical bearer of the semantic relation – that is, the relation is inalienable from its semantics regardless of its contextual realisations and is said to be its extra-sentential semantic role (Lakova 2015). In this study we aim at describing the set of nouns that may enter in a predicate–argument or predicate–adjunct relation with a given verb on the basis of purely semantic criteria, thus including morphosemantic relations but going beyond them.

The description of the semantic relations based on the conceptual structure of the predicates includes the following elements:

(i) A **definition of the relation**, which is elaborated after an analysis of the definitions of the corresponding frame element across the frames it is found and/or an analysis of the scope of its counterpart semantic role in VerbNet (if relevant). The definition aims at providing maximum generalisation of the semantic properties of the frame element. For instance, the FE *Agent* is found in 180 frames – by studying the 180 corresponding definitions, we identify and confirm the relevant components of the description of the FE that may be used in the formulation of the relation, such as: *individual*, *sentient*, *animate, conscious entity, volitional, intentionally.*

(ii) **Relationships with more abstract and/or more specific relations** are formulated on the basis of the extracted relations of inheritance between pairs of frames and corresponding FEs in these pairs. In defining a more abstract relation we take into consideration its correspondence with a semantic role in VerbNet if we consider it relevant: for instance, we did so with respect to the frame elements discussed below in this subsection, but we ignored the correspondence between *Cognizer* (FrameNet) and *Agent* (VerbNet) with verbs such as *decide*, *choose*, as well as between *Cognizer* (FrameNet) and *Experiencer* (VerbNet) with verbs such as *discover*, *encounter*, *understand* and decided to posit it as a separate relation on the grounds of its distinct semantics – as part of the conceptual description of verbs of cognition (see the discussion on the semantic relation *Cognizer* in Section 6.14).

The more specialised variants of a relation which are not defined as separate relations correspond to FrameNet frame elements and serve as illustrations of the relation. The relationship between all the levels of generalisation is preserved and one may lump together specific relations or may employ more detailed ones. For instance, the *Agent* relation represents the most abstract level with a corresponding *Agent* semantic role in VerbNet. On the basis of the relations of inheritance between frames and between frame elements, it is further specified by the relations *Communicator* (*Agent.Communicator*), *Self_mover* (*Agent.Self_mover*), *Ingestor* (*Agent.Ingestor*), *Creator* (*Agent.Creator*), among others. *Communicator* is further specified as *Arguer, Informer, Claimant*[24]. Thus, the following hierarchy is established: *Agent* > *Communicator* > *Arguer* | *Informer* | *Claimant*. In this case, we have proposed *Communicator, Self_mover, Ingestor, Creator* as new relations (representing variants of the relation *Agent*) as they describe differences between various *Agents* which correspond to the the semantic properties of the verb classes for which they are defined – verbs of communication, verbs of motion, verbs of consumption, among others, and are sufficiently inclusive. On the basis of the inheritance, we are also able to define further specialisations of the relation that we currently consider too fine-grained, such as *Arguer* (*Agent.Communicator.Arguer*), *Informer* (*Agent.Communicator.Informer*), among others.

The obtained hierarchy is largely based on the observed linguistic data, but, where needed, we have either supplemented them or introduced changes. In some cases the internal organisation of FrameNet does not provide information about inheritance. In such instances we rely on the mapping between VerbNet semantic roles and the WordNet structure to define such a relation. This was done, for instance, for the frame element *Competitor*, which is not directly related through inheritance to *Agent*, but after analysing its VerbNet counterpart (*Agent*) and its semantics, we posited it as a subtype of *Agent* and defined the respective relation.

(iii) The **selectional preferences characterising each relation** are based on the semantic features derived from the selectional restrictions in VerbNet and the semantic types in FrameNet. They serve as a point of departure for defining the relevant semantic classes or subclasses of nouns to which verbs from particular classes are linked through a given relation. For instance, for the nouns entering in the relation *Agent*, we assign the features associated with the particular verbs, e.g. [Sentient] from FrameNet, [+ANIMATE], [+ORGANIZATION], [+MACHINE], etc. from VerbNet. These features are then analysed and translated into corresponding classes of noun synsets representing the default selectional preferences which may

[24] The frame elements and their definitions in the frames to which they belong may be viewed at: https://www.clres.com/clr/fetax.php.

be further specified for different subclasses of verbs (synsets assigned a particular frame) or even for a given synset; compare Example 1a and 1b where the selectional preferences on the nouns entering in the relation *Agent.Communicator* are different for the verbs evoking the frames Communication and Gesture. The general assumptions underlying the definition of selectional preferences are outlined in Section 7. A more detailed representation of the preferences formulated for the relation *Theme* is laid out in Section 8, Table 1.

Example 1. Selectional preferences for the relation *Agent.Communicator* in different frames.

(1a) **[The company | director]**$_{\text{AGENT.COMMUNICATOR}}$ **indicated** that there would be cutbacks. (persons, organisations and social groups)

Frame: Communication, Relation: Agent.Communicator, Selectional preferences: Volitional_human_entity (persons, organisations, social groups)

(1b) **[Mark]**$_{\text{AGENT.COMMUNICATOR}}$ **nodded** to confirm. (predominantly persons)

Frame: Gesture, Relation: Agent.Communicator, Selectional preferences: Volitional_human (persons)

6. Defining the semantic relations

Below we define the main inventory of semantic relations. As noted above, as we preserve the hierarchy of the respective frame elements on which we found the relations, we are able to define deeper-level subrelations (if needed) or we may resort to a more general level of description. In fact, for most of the relations we have defined subtypes which capture semantic distinctions within the relation but are not (and cannot be) exhaustive. The more general instances which do not fit in any of the defined subtypes are assigned the parent relation, e.g. *Agent*. The relation often bears the name of the frame element to which it links the predicate (or the name of a more general or a more specific FE); the two meanings are distinguished by specifying explicitly whether it is the relation or the participant that is under discussion.

6.1. Agent[25]

The relation *Agent* is defined between a predicate and an animate or a quasi-animate (a social group, an organisation or the like) participant in its conceptual structure, typically a person or another sentient entity, who consciously or volition-

[25] Sections 6.1. – 6.15. represent a revised and extended version of the classification proposed in Leseva et al. 2018: Лесева, Св., Ив. Стоянова, Хр. Кукова, М. Тодорова. Интегриране на субкатегоризационна информация в релационната структура на УърдНет. – *Български език*, 65, 2, 2018, с. 11 – 40. (Leseva, S., I. Stoyanova, H. Kukova, M. Todorova. Integrirane na subkategorizatsionna informatsiya v relatsionnata struktura na Uardnet. – *Balgarski ezik*, 65, 2, 2018, pp. 11 – 40.)

ally initiates and carries out an act, activity or the like, and has control over it, or who produces a particular effect, change, etc. (in another participant); the *Agent* participant exists independently of the situation. This relation is defined for verbs belonging to almost all semantic primitives (to the exception of verbs of weather) to the extent that even situations described by verbs that are typically not considered agentive may be brought about by an animate volitional cause: consider the examples of verbs of emotion (Example 2b), perception verbs (Example 2c), cognition verbs (Example 2d).

On the basis of the analysis of the inheritance between the semantic role/FE *Agent* and more specific agentive FE realisations we defined several non-exhaustive subtypes of the relation[26]:

(1) *Agent.Communicator* – an animate or a quasi-animate entity that creates and/or conveys a message or a piece of information using a particular medium of communication (Example 2d). More specific variants include (non-exhaustively): *Arguer, Claimant, Complainer, Informer, Interlocutor, Questioner, Speaker*.

(2) *Agent.Creator* – an animate or a quasi-animate entity that makes, constructs, creates, produces, manufactures a material or non-material product or brings about the existence of some entity (Example 2e). Specific variants include: *Author, Artist, Producer*.

(3) *Agent.Donor* – a person, a group or a quasi-animate entity that begins in the possession of some object, asset or non-material property and causes it to pass into the possession of some other person, group or quasi-animate entity (Example 2f). Specific variants include: *Offerer, Submittor, Supplier, Surrenderer*.

(4) *Agent.Ingestor* – an animate entity that consumes food, drink or another substance (Example 2g).

(5) *Agent.Competitor* – an animate or a quasi-animate entity that takes active participation in some kind of competition, usually with the purpose of beating an opponent (Example 2h).

(6) *Agent.Self_mover* – an animate entity (as well as a vehicle or another mobile machine under the control of a person) that moves in space along a path under its own power and by means of its own body (Example 2i).

(7) *Agent.Changer* – an animate or a quasi-animate entity that brings about some change in the state, condition, form, structure, integrity or another essential attribute of the object it acts upon (Example 2j). Specific variants include: *Deformer, Destroyer, Healer, Killer*.

[26] Instead of the relation we describe the second member of the relation (corresponding to a participant or another element in the conceptual description of the predicate).

Example 2. The relation *Agent*.

(2a) **[The bartender]**$_{\text{AGENT}}$ **shook** the shaker energetically.

(2b) **[He] amused**$_{\text{AGENT}}$ the children the whole evening.

(2c) **[The aspiring artist]**$_{\text{AGENT}}$ **showed** his last painting.

(2d) **[The lawyer]**$_{\text{AGENT.COMMUNICATOR}}$ easily **refuted** the charges.

(2e) **[The renowned novelist]**$_{\text{AGENT.CREATOR}}$ **wrote** a sequel to his bestseller.

(2f) **[The firm]**$_{\text{AGENT.DONOR}}$ **donated** part of their profit to a local charity.

(2g) **[They]**$_{\text{AGENT.INGESTOR}}$ **gobbled down** the cake.

(2h) **[The champion]**$_{\text{AGENT.COMPETITOR}}$ **beat** his opponent in the first round.

(2i) **[She]**$_{\text{AGENT.SELF\_MOVER}}$ **walked** along the canal.

(2j) **[The workers]**$_{\text{AGENT.CHANGER}}$ **destroyed** the old bridge.

6.2. Cause

The relation *Cause* is defined between a predicate and either an inanimate participant, such as a natural phenomenon, an event, a process, a situation or the like, or less typically – an animate or a quasi-animate entity (Example 3c) – that initiates and brings about an effect, change, etc., but does not act intentionally, consciously or volitionally. These semantic features of the *Cause* distinguish it from the *Agent* with which it can often alternate, e.g.: *The boy broke the window* vs. *The wind broke the window*. The following subtypes are defined:

(1) *Cause.Effector* – a physical or an abstract inanimate cause or an animate cause that brings about an effect or some change in another participant[27] (Example 3a, b, c). Specific variants include: *Grounding_cause, Preventing_cause*.

(2) *Cause.Stimulus* – any phenomenon, situation, mental content or state, etc. that causes an emotional or a psychological response in a participant, such as an *Experiencer* (Example 3d).

Example 3. The relation *Cause*.

(3a) **[The wind]**$_{\text{CAUSE.EFFECTOR}}$ **blew off** several tiles from the roof.

(3b) **[The price raise]**$_{\text{CAUSE.EFFECTOR}}$ **led** to a full-blown crisis.

(3c) **[Even one extra member]**$_{\text{CAUSE.EFFECTOR}}$ **will change** the ratio in the group.

(3d) **[The illness]**$_{\text{CAUSE.STIMULUS}}$ **tormented** her for several years.

---

[27] Cf. Van Valin and LaPolla (Van Valin, LaPolla 1997) on the distinction between *Agents*, *Effectors* and *Forces*; the name Effector is borrowed from their work, but applied more broadly.

6.3. Patient

The relation *Patient* is defined between a predicate and a participant in its conceptual structure that is directly affected by or acted upon or experiences some change or influence (by an *Agent* or another participant) and as a result undergoes a change of state, condition, form, structure, integrity, etc. The relevant participant exists independently of the situation.

(1) *Patient.Body_part* – the part of the body of a participant in the situation that undergoes an injury or another change or is otherwise affected (Example 4d).

(2) *Patient.Victim* – an animate entity (or a group of such entities) that suffers an injury, violence or other harmful influence (Example 4e). Specific variants include: *Victim, Executed*.

(3) *Patient.Object_of_physical_impact* – an inanimate entity that undergoes physical influence (contact, impact, etc.) with or without an actual change in its state, condition, form, structure, function, etc. (Example 4f). Specific variants include: *Target, Impactee*.

(4) *Patient.Ingested_substance* – substances or products, such as food, drink, medications, etc. that an animate entity (*Ingestor*) takes into his or her organism (Example 4g). Specific variants include: *Ingested_substance, Food*.

Example 4. The relation *Patient.*

(4a) The peasants **ground [the wheat]**$_{\text{PATIENT}}$.

(4b) The shaman **healed [the wounded warrior]**$_{\text{PATIENT}}$.

(4c) The workers **blasted [the old building]**$_{\text{PATIENT}}$.

(4d) He **hit [his head]**$_{\text{PATIENT.BODY\_PART}}$ in the wall.

(4e) **[He]**$_{\text{PATIENT.VICTIM}}$ **was injured** in an accident.

(4f) The archer **hit [the target]**$_{\text{PATIENT.OBJECT\_OF\_PHYSYCAL\_IMPACT}}$, just missing the bull's eye.

(4g) The children **ate [the cake]**$_{\text{PATIENT.INGESTED\_SUBSTANCE}}$.

6.4. Experiencer

The relation *Experiencer* relates a predicate and an animate participant in its conceptual structure that experiences or displays a feeling, an emotional or psychological state or reaction or a physical perception or state. Verb-wise the relation is most typical for verbs of emotions, verbs of perception related to the functions of the body or the like, etc. We distinguish two subtypes of the relation which are defined as follows in terms of the noun members involved:

(1) *Experiencer.Emotion* – a sentient animate being, usually a person, that experiences a feeling, an emotional or psychological state or reaction (Example 5a).

(2) *Experiencer.Sensation* – an animate being that undergoes a physical sensation (Example 5b).

Example 5. The relation *Experiencer*.

(5a) She **liked [the child]**$_{\mathrm{EXPERIENCER.EMOTION}}$.

(5b) **[The children]**$_{\mathrm{EXPERIENCER.SENSATION}}$ **basked** in the warm sun.

6.5. Perceiver

The relation *Perceiver* is defined between a predicate and an animate or a quasi-animate participant in its conceptual structure that perceives objects or phenomena through one of the senses. In the traditional theory of semantic roles *Perceivers* are usually subsumed under the scope of *Experiencer*. On the basis of the well-known dichotomy between the type of perception with respect to the degree of volitional and conscious activity on the part of the subject of perception, we distinguish between two subtypes: *Agentive_perceiver* and *Passive_perceiver*.

(1) *Agentive_perceiver* – a sentient animate entity that consciously or volitionally perceives objects or phenomena through one of its senses (Example 6a).

(2) *Passive_perceiver* – a sentient animate entity that perceives objects or phenomena through one of the senses not necessarily on purpose and without conscious control (Example 6b).

Example 6. The relation *Perceiver*.

(6a) **[The guests]**$_{\mathrm{PERCEIVER.AGENTIVE\_PERCEIVER}}$ **observed** the eclipse from the balcony.

(6b) **[The sleeping dog]**$_{\mathrm{PERCEIVER.PASSIVE\_PERCEIVER}}$ **smelled** the food.

6.6. Recipient

The relation *Recipient* is defined between a predicate and a sentient animate or quasi-animate participant in its conceptual structure that ends up in the possession of a concrete or an abstract entity, usually a *Theme*, as a result of the situation described. This relation is most typical of verbs of (transfer of) possession. More specific variants include: *Addressee, Buyer, Seller*.

Example 7. The relation *Recipient.*

(7a) She **paid** 100 million **[to the kidnapper]**$_{\mathrm{RECIPIENT}}$.

(7b) The movie star **left** his wealth **[to an animal rescue organisation]**$_{\mathrm{RECIPIENT}}$.

6.7. Theme

The relation *Theme* is defined between a predicate and an animate, a quasi-animate or an inanimate participant in its conceptual structure that is acted upon or affected in some way in the course of the evolving of the situation, especially by changing its location, and/or by being characterised or evaluated in terms of its state, position, condition, properties, etc. The *Theme* participant does not have control over the situation and does not undergo changes in its structure, form, function or essential properties (unlike participants involved in the relation *Patient*). The relation encompasses a broad range of more specific variants, such as *Affected_entity, Evaluee, Sought_entity, Chosen, Escapee, Suspect, Examinee, Prisoner*, which are subsumed under the general relation (Example 8a, b). Below are presented the coherent subtypes we have identified:

(1) *Theme.Situation* – an act, activity, event, phenomenon that is being carried out, performed, etc. in the course of the situation described or is acted upon or affected or characterised in some way (Example 8c). More specific variants include: *Phenomenon, Act, Activity, Event*.

(2) *Theme.Information_content* – information or cognitive content that is being conveyed or exchanged in the course of the situation (Example 8d, e). Specific variants include: *Documents, Mental_content, Message, Opinion*.

(3) *Theme.Emitted_entity* – a substance or material or another physical entity or property (such as energy) that is emitted from a *Source*, possibly as a result of the action of a *Cause.Effector* or an *Agent* (Example 8f). Specific variants include: *Light, Precipitation, Radiation, Secretion, Sound*.

(4) *Theme.Transferred_possession* – a physical or an abstract entity that is the object of exchange or transfer of possession (Example 8g). Specific variants include: *Money, Assets, Goods, Possession*.

(5) *Theme.Body_part* – a part of the body of an *Agent* who affects or acts upon it especially by changing its position or spatial configuration with respect to the remaining part of the body; the body part does not undergo any change in its condition, shape, structure, function, etc. (Example 8h).

Example 8. The relation *Theme*.

(8a) The council **approved [us]**$_{\text{THEME}}$ for the job.

(8b) The child **rolled [the ball]**$_{\text{THEME}}$ to his brother.

(8c) The scientists **detected [a strange gamma ray burst]**$_{\text{THEME.SITUATION}}$.

(8d) They **shared [some very interesting news]**$_{\text{THEME.INFORMATION\_CONTENT}}$ with us.

(8e) We soon **realised [our mistake]**$_{\text{THEME.INFORMATION\_CONTENT}}$.

(8f) If in danger, these animals **secrete [poison]**$_{\text{THEME.EMITTED\_ENTITY}}$.

(8g) They **sold [the old flat]**$_{\text{THEME.TRANSFERRED\_POSSESSION}}$ and **bought [a new one]** $_{\text{THEME.TRANSFERRED\_POSSESSION}}$.

(8h) The horse **lifted [his head]**$_{\text{THEME.BODY\_PART}}$.

6.8. Means

The relation *Means* is defined between a predicate and a physical or an abstract participant in its conceptual structure that is used in carrying out the situation or which mediates or facilitates the situation. The following subtypes of the relation are further defined:

(1) *Means.Instrument* – an implement, a device or another object that is manipulated by an *Agent* or another participant in order to carry out the action, activity, event, etc. (Example 9a). Specific variants include: *Instrument, Weapon, Heating_instrument, Connector*.

(2) *Means.Mediator* – a physical object, a substance or an abstract entity that is employed by an *Agent* or another participant and which indirectly enables this participant to carry out the action, activity, event, etc., thus fulfilling its (natural) purpose (Example 9b, c). More specific variants include *Medication, Means, Vehicle*.

(3) *Means.Manner* – a quality, a skill, an ability, an action, etc. that enables an *Agent* or another participant to carry out the action, activity or event (Example 9d).

(4) *Means.Medium* – a technological device or a physical or an abstract entity that is used by a *Communicator* or another agentive participant in the process of communication as a channel or medium of conveying a message to the *Addressee* (Example 9e).

Example 9. The relation *Means.*

(9a) The soldier **took aim [with his bow]**$_{\text{MEANS.INSTRUMENT}}$.

(9b) Many doctors **treat** the common cold **[with antibiotics]**$_{\text{MEANS.MEDIATOR}}$.

(9c) **[His words]**$_{\text{MEANS.MEDIATOR}}$ **confirmed** my suspicions.

(9d) **[Memorising]**$_{\text{MEANS.MANNER}}$ **does not mean acquiring** true knowledge.

(9e) He **told** us the news **[over the phone]**$_{\text{MEANS.MEDIUM}}$

6.9. Result

The relation *Result* is defined between a predicate and a participant or an element in its conceptual structure that denotes the concrete or abstract result, end product, state or effect occurring as a direct consequence of the situation described by the verb. The relation *Result* has the following subtypes:

(1) *Result.Final_state* – the state of a participant occurring as a direct consequence of its undergoing a change or experiencing another effect in the course of the situation (Example 10a).

(2) *Result.Final_category* – the category which the participant ends in as a direct consequence of its involvement in the situation described by the verb (Example 10b).

(3) *Result.Final_value* – the value that an *Attribute* of a participant assumes as a consequence of the participant's involvement in the situation (Example 10c).

(4) *Result.Created_entity* – a concrete or an abstract entity which comes into existence as a direct consequence of the situation described (Example 10d, e). More specific variants include *Idea, Copy, Image, Product*.

(5) *Result.Effect* – a change or a state-of-affairs which occurs as a direct consequence of the situation (Example 10f, g).

Example 10. The relation *Result.*

(10a) The prices **rose [to unprecedented levels]**$_{\text{RESULT.FINAL\_STATE}}$.

(10b) He **turned** into **[a devout follower]**$_{\text{RESULT.FINAL\_CATEGORY}}$.

(10c) The bank's interest **was reduced [to 0.1%]**$_{\text{RESULT.FINAL\_VALUE}}$.

(10d) The kids **invented [a new game]**$_{\text{RESULT.CREATED\_ENTITY}}$.

(10e) The company **manufactured [wooden toys]**$_{\text{RESULT.CREATED\_ENTITY}}$.

(10f) The new enforced rules **caused [riots]**$_{\text{RESULT.EFFECT}}$.

(10g) The spread of the virus **prompted [a lockdown]**$_{\text{RESULT.EFFECT}}$.

6.10. Destination

The relation *Destination* is defined between a predicate and an element in its conceptual structure that expresses a concrete or an abstract place or point which a participant, usually an *Agent* or a *Theme*, is aiming for or wants to reach in the course of the situation. We posit two subtypes of the relation, which are defined as follows:

(1) *Destination.End_point* – the point in space which is reached by a participant (usually an *Agent* or a *Theme*) as a result of the situation described by the predicate (Example 11a, b, c).

(2) *Destination.Goal* – a desirable state-of-affairs which a participant (usually an *Agent*) wants to achieve through the situation described by the verb (Example 11d, e).

Example 11. The relation *Destination.*

(11a) Tomorrow we **will fly [to Barcelona]**$_{\text{DESTINATION.END\_POINT}}$.

(11b) The ball **rolled [into the bushes]**$_{\text{DESTINATION.END\_POINT}}$.

(11c) She **turned** her attention **[to Mark]**$_{\text{DESTINATION.END\_POINT}}$.

(11d) He **will attempt [to prove this theorem]**$_{\text{DESTINATION.GOAL}}$.

(11e) He **failed [to get his degree]**$_{\text{DESTINATION.GOAL}}$.

6.11. Source

The relation *Source* is defined between a predicate and an element in its conceptual structure which expresses a place or a point (either a concrete or an abstract one) where the situation arises or from which a participant begins to move or act. It is further divided into the following subrelations:

(1) *Source.Starting_point* – a concrete or an abstract place, point, or object in space where the situation begins (Example 12a, b).

(2) *Source.Emitter* – a place or an object from which the emission of light, sound, etc. is discharged (Example 12c).

(3) *Source.Initial_state* – the initial, previous or original state of a participant which changes in the course of the situation (Example 12d).

(4) *Source.Initial_category* – a category to which a participant belonged which changes in the course of the situation (Example 12e).

(5) *Source.Initial_value* – the value that an attribute of a participant had which changes in the course of the situation (Example 12f).

(6) *Source.Material* – a physical entity, especially raw material, a substance, an ingredient or the like which is transformed into another substance, object or end product in the course of the situation (Example 12g).

Example 12. Relation *Source.*

(12a) Tomorrow we **fly [from Sofia]**$_{\text{SOURCE.STARTING\_POINT}}$ to Varna.

(12b) The criminals **stole** 100 thousand leva **[from the bank]**$_{\text{SOURCE.STARTING\_POINT}}$.

(12c) **[The beacon]**$_{\text{SOURCE.EMITTER}}$ **emits** bright pulsating light.

(12d) He **pulled** her **[from her lying position]**$_{\text{SOURCE.INITIAL\_STATE}}$ and forced her to stand.

(12e) She **improved [from a mediocre piano player]**$_{\text{SOURCE.INITIAL\_CATEGORY}}$ to a true proficient.

(12f) The unemployment among young people **has fallen [from 25%]**$_{\text{SOURCE.INITIAL\_VALUE}}$ to 17%.

(12g) They **built** the house **[out of stone]**$_{\text{SOURCE.MATERIAL}}$.

6.12. Location

The relation *Location* is defined between a predicate and an element in its conceptual structure which represents a concrete or an abstract place or region where the situation occurs. The relation encompasses the following subtypes:

(1) *Location.Place* – the point in space where the situation takes place (Example 13a).

(2) *Location.Area* – a region where the situation takes place (Example 13b).

(3) *Location.Container* – an object with a specific capacity where the situation takes place, especially where something is placed or positioned (Example 13c).

(4) *Location.Body_location* – a place on or inside the body of a participant (e.g., an *Experiencer*, *Patient*, etc.) who/which is involved in the situation (Example 13d).

(5) *Location.Path* – the trajectory of motion (between the *Source* and the *Destination*) which a participant moves along (Example 13e).

Example 13. The relation *Location.*

(13a) I have **lived [in Sofia]**$_{\text{LOCATION.PLACE}}$ for 20 years.

(13b) Heavy rains **are expected [across the North-East of the country]**$_{\text{LOCATION.AREA}}$.

(13c) I **went through [the whole wardrobe]**$_{\text{LOCATION.CONTAINER}}$.

(13d) The kid **wore** a blue hat **[on his head]**$_{\text{LOCATION.BODY\_LOCATION}}$.

(13e) I **walked [along the river]**$_{\text{LOCATION.PATH}}$.

6.13. Protagonist

The relation *Protagonist* is defined between a predicate and an animate or a quasi-animate entity which is the main participant and is directly involved in the situation. The relation expresses mostly predicate-argument relations of mixed typology, combining characteristics of *Agent* and *Patient* (in the setting of semantic roles participants with such typology are usually assigned the role *Theme*), as often the *Protagonist* participant is involved in activities or actions in which he or she both acts upon and is acted upon and/or undergoes the effect or change of state, condition, position, etc. that takes place in the course of the situation (Example 14a, b, c). The *Protagonist* participant may also undergo states or processes in which he or she has at least partial control (Example 14d, f) or may experience internally evolving states and processes (Example 14e); the noun member of the *Protagonist* relation may also be an animate entity who has certain characteristics or properties or behaviour or participates in some relationship (Example 14g, h).

Example 14. The relation *Protagonist.*

(14a) **[The girl]**$_{\text{PROTAGONIST}}$ **sat down** quietly.

(14b) **[He]**$_{\text{PROTAGONIST}}$ **came back** from the dead.

(14c) **[The child]**$_{\text{PROTAGONIST}}$ **woke up** and jumped from the bed.

(14d) **[My mother]**$_{\text{PROTAGONIST}}$ **freaked out** when she saw the broken vase.

(14e) **[He]**$_{\text{PROTAGONIST}}$ **died** a violent death.

(14f) **[The cat]**$_{\text{PROTAGONIST}}$ **waited** for him to open the can.

(14g) **[The minister]**$_{\text{PROTAGONIST}}$ **went back** on his promise to build a hospital.

(14h) **[They]**$_{\text{PROTAGONIST}}$ **comply** with the law.

6.14. Cognizer

The relation *Cognizer* is defined between a predicate and an animate or a quasi-animate entity, in most cases a person, who: (a) performs a cognitive activity (observing, thinking, reasoning, evaluating, decision-making, etc.) in order to determine or to discover a particular fact, knowledge, information, to form an opinion or an attitude, to make a decision, etc.; or (b) is the subject of a mental or cognitive process or state. Although the semantic role *Cognizer*, on which the relation is based, is considered as a kind of *Experiencer* in some accounts (Van Valin, LaPolla 1997: 85), the entities assigned the frame element *Cognizer* in FrameNet have distinct semantics and express agentive involvement in cognitive activities or processes (Example 15a–e) unlike true *Experiencers*; this prompted us to posit *Cognizer* as a separate relation. Specific variants include: *Assessor*, *Audience*, *Believer*, *Inspector, Researcher, Reader*.

Example 15. The relation *Cognizer.*

(15a) **[He]**$_{\text{COGNIZER}}$ **considered** both options.

(15b) **[Philosophers]**$_{\text{COGNIZER}}$ **have contemplated** on the meaning of life for ages.

(15c) **[She]**$_{\text{COGNIZER}}$ **chose** the lesser evil.

(15d) **[The girl]**$_{\text{COGNIZER}}$ **believed** she was right.

(15e) **[Mark]**$_{\text{COGNIZER}}$ **expects** to win the competition.

6.15. Property

The relation *Property* is defined between a predicate and an element in its conceptual structure that expresses an attribute, a quality or a characteristic assigned to a participant (such as a *Theme* or a *Patient*) involved in the situation described by the verb. The relation is further divided into the following subtypes:

(1) *Property.Attribute* – a property such as *Quantity, Number, Value, Age, Size, Dimension, Temperature, Speed* whose value changes or which is subjected to observation, evaluation, etc. (Example 16a, b).

(2) *Property.Role* – a function or a role adopted by a participant in the situation (Example 16c, d).

Example 16. The relation *Property.*

(16a) The garden **smells [of fresh herbs]**$_{\text{PROPERTY.ATTRIBUTE}}$.

(16b) **[Electricity prices]**$_{\text{PROPERTY.ATTRIBUTE}}$ **will increase** next year.

(16c) They **appointed** him **[as captain of the team]**$_{\text{PROPERTY.ROLE}}$.

(16d) They **trained** him **[as a professional swimmer]**$_{\text{PROPERTY.ROLE}}$ from an early age.

7. Modelling semantic restrictions

As noted above, in order to define the selectional preferences that verbs impose on the main participants in their conceptual structure, we use the combined information encoded in the three resources, in particular WordNet semantic primitives, FrameNet semantic types and VerbNet selectional restrictions.

7.1. General prerequisites

The generalisations proposed in the resources under discussion are either similar or complementary to each other: for instance, the semantic type Sentient in FrameNet roughly corresponds to the feature [+ANIMATE] in VerbNet and the set union of the WordNet synsets with the primitives noun.animal (animals) and noun.person (persons). Further refinements may be needed: for example, to exclude non-relevant members of these classes, or to include quasi-animate entities (e.g. companies or other organisations) which are semantically licensed by certain classes of verbs. In addition, VerbNet proposes the feature [+INT_CONTROL] (internal/intentional control), thus providing a further narrowing down of the Sentient semantic type which translates as the notion of a volitional *Agent*, a restriction relevant for strictly agentive verbs, such as {*beat:2, beat up:1, work over:1*} ('give a beating to; subject to a beating, either as a punishment or as an act of aggression') or {*persecute:1, oppress:2*} ('cause to suffer'). In a like manner, we explored the restrictions encoded in the three resources with a view to adopting a consistent unified set.

In addition, the observations point to the necessity for imposing even further restrictions that may distinguish between different subtypes of a relation. For instance, while [+INT_CONTROL] accounts for volitionality, different classes of verbs require narrower specifications. Some groups of verbs, such as *build* and *run* may take as *Agents* both persons and animals, while others, such as *study* and *dress* are restricted to human beings alone; further, as noted above, there are verbs such as *judge* and *hire* which select for human entities (including organisations in addition to persons). This latter semantic feature corresponds to the VerbNet selectional restriction [+ORGANIZATION]. To account for such finer distinctions, we either adopted existing sets of restrictions, such as the VerbNet [+ANIMATE | +ORGANIZATION] or formulated new ones, e.g. by differentiating between [+ANIMATE] (Sentient in FrameNet's terms) and strictly Human.

Even more fine-grained semantic distinctions stem from the analysis of individual predicates or groups of predicates in a given frame or group of frames. Consider the example of verbs related to drying, which evoke the Cause_to_dry frame in FrameNet. Some of them, e.g. {*dry:1*} 'remove the moisture from and make dry', have very general selectional preferences for the thing being dried (the *Dryee*), in this case – a physical entity or a surface; others such as {*drip-dry:1*} 'dry by hanging up wet', {*spin-dry:1*} 'dry (clothes) by spinning and making use of centrifugal forces' and {*tumble dry:1*} 'dry by spinning with hot air inside a cylinder' impose the restriction Garment (clothes and textiles), and yet others, such as {*blow-dry:1*} 'dry hair with a hair dryer' have the very specific preference Hair (Leseva et al. 2020).

Towards making up for the different granularity and scope of the semantic features that underlie the selectional preferences, we complement the semantic primitives in WordNet, the VerbNet semantic roles and restrictions and the FrameNet semantic types with additional semantic information which facilitates the formulation of fine-grained preferences as in the 'dry' examples above; a source of such information are the definitions of particular frame elements in FrameNet. One notable example is the restriction Comestible (types of food and drink) corresponding to the definition of the frame element *Ingestibles* in the frame Ingestion (evoked by verbs such as *eat*, *drink*, *breakfast*, *dine*), which is defined as 'the entities that are being consumed', in particular types of food and drink. As this preference is very specific for the verbs of consumption, we posit it as a separate one (see also Example 17).

The so-formulated inventory includes selectional preferences of different granularity and is open to expanding as well as to revising the already defined ones. In analysing and validating the data, we start from exploring the more general restrictions and then proceed to defining more fine-grained ones, where necessary, on the basis of the available semantic information and observations.

The semantic restrictions imposed on the verb's arguments often align with a particular subtree or subtrees of noun synsets which correspond to salient ontological classes or subclasses (Agirre, Martinez 2002; Koeva 2010; Di Fabio et al. 2019). In Section 7.2 we present the main types of selectional preferences imposed by verbs on the nouns they combine with, which we have formalised by means of classes of noun synsets. The actual formal restrictions illustrated in Example 17 will be presented in detail in Section 8 for the relation *Theme*.

7.2. Specification of general preferences[28]

Below we present an outline of the main types of selectional preferences based on the analysis of WordNet synsets with manually verified assigned frames (cf. Section 4). We determine 8 major classes: Animate_and_groups; Concrete; Content; Abstract; Spatial_description; Temporal_description; Eventuality; Attribute. These names are in fact shorthands for the formalised restrictions. For instance, the restriction Comestible may be translated as the union (∪) of three classes of synsets: the trees stemming from two synsets denoting foods and drinks and the synsets belonging to the respective semantic primitive (Example 17). In combining the primitive with the tree structure, we seek to achieve exhaustiveness of the possible nouns that meet these selectional preferences.

Example 17. Formal description of the restriction Comestible applying to the FE *Ingestible* in the FN frame Ingestion:

{*food:2; solid food:1*} ∪ {*food:1; nutrient:1*} ∪ noun.food

where:

(a) {*food:2; solid food:1*} 'any solid substance (as opposed to liquid) that is used as a source of nourishment';

[28] The preferences laid out in Section 7.2 represent a revised and extended version of the classification motivated and presented in Leseva et al. 2019: Leseva, S, I. Stoyanova, M. Todorova, H. Kukova. A Theoretical Overview of Conceptual Frames and Semantic Restrictions on Frame Elements. – *Балканско езикознание/Linguistique Balkanique*, LVIII, 2, 2019: pp. 172 – 186.

(b) {*food:1; nutrient:1*} 'any substance that can be metabolised by an animal to give energy and build tissue';

(c) noun.food – the semantic primitive assigned to synsets denoting types of nutrients, foods and drinks; not necessarily disjoint with (a) and (b).

As the classes of preferences presented below are generalisations made from the observations on particular classes of verbs, they are non-exhaustive and are subject to further extension and refinement in the process of studying other classes of predicates. Such finer distinctions, e.g. *Concrete > Physical_entity > Astronomical_entity, Concrete > Physical_object > Projectile, Concrete > Fluid > Precipitation, Concrete > Fluid > Excreta, Content > Documents*, among others, are proposed in Section 8 for the relation *Theme*. Each general restriction is illustrated by a frame element (in a given frame) to which it is applicable.

(1) Animate entities and groups

(1.1) Sentient_entity: encompasses sentient (groups of) beings, including quasi-animate entities (organisations and social groups) (Giving:Recipient[29]);

(1.2) Animate_entity: a subset of the class of Sentient_entities which includes persons and animals but excludes quasi-animate entities (Cause_harm:Victim);

(1.3) Volitional_entity: refers to persons, animals and quasi-animate entities as well as to automated self-operating devices performing controlled (most often human-like) activities or actions; Volitional_entities have (intentional) control of the performed activity, action, process or event and act consciously or volitionally or in a like manner (Building:Agent);

(1.4) Volitional_animate: a subset of the class of Volitional_entities which includes persons and animals but excludes quasi-animate entities and automated self-operating devices; applies to volitional activities of living beings (Ingestion:Ingestor);

(1.5) Volitional_animate_machines: a subset of the class of Volitional_entities which includes Volitional_animate as well as automated self-operating devices (Self_motion:Self_mover);

(1.6) Volitional_human: a subset of the class of Volitional_animate which excludes non-human entities, i.e. animals (Forming_relationships:Partners);

(1.7) Volitional_human_entity: a subset of the class of Volitional_entities which includes Volitional_human along with organisations and social groups (Hiring:Employer);

(1.8) Volitional_human_machines: a subset of the class of Volitional_entities which includes Volitional_human as well as automated self-operating devices (Create_physical_artwork:Creator);

[29] The format for the examples is (Frame:Frame Element).

(1.9) Human_entity: a subset of the class of Sentient_entities which includes persons as well as quasi-animate entities but unlike Volitional_human_entities does not require conscious or volitional activity (Judgment:Evaluee);

(1.10) Human: a subset of the class of Animate_entities which includes persons but unlike Volitional_human does not require conscious or volitional activity (Imprisonment:Prisoner);

(1.11) Body_part: although not strictly animate, Body_part refers to parts of the body of animate beings which are involved in the activity, action, process or event (Cause_harm:Body_part).

(2) Concrete

(2.1) Physical_entity: includes physical objects and substances but excludes abstract entities, eventualities (activities, actions, events, processes, etc.), attributes and the like (Filling:Theme);

(2.2) Physical_object: a subset of the class of Physical_entities which specifically excludes substances, as well as locations, natural objects, geological formations, etc. that are construable as locations (Apply_heat:Heating_instrument);

(2.3) Artifact: a subset of the class of Physical_objects which is restricted to man-made entities (Intentionally_create:Created_entity);

(2.4) Substance: a subset of the class of Physical_entities which refers to substances regardless of their physical form or phase (Change_of_phase:Patient);

(2.5) Fluid: a subset of the class of Substances which refers to fluids and gases (Cause_fluidic_motion:Fluid);

(2.6) Solid: a subset of the class of Physical_objects which includes solid bodies of certain shape (Cause_to_fragment:Whole_patient);

(2.7) Vehicle: a subset of the class of Artifacts that includes means of transportation, both motorised and non-motorised (Operate_vehicle:Vehicle);

(2.8) Comestible: a subset of the class of Physical_entities which refers to substances and products that can be consumed (Tasting:Food);

(2.9) Garment: a subset of the class of Artifacts which includes clothing, textiles and the like (Wearing:Clothing);

(2.10) Emission: a subset of the class of Physical_entities which refers to substances, physical entities or properties, such as sound, light, waves, smell, energy, etc. that are discharged, usually from a *Source* (Emitting:Emission);

(2.11) Sound: a subset of the class of Emission which refers to sounds, noises and other acoustic perceptions (Make_noise:Sound);

(2.12) Light: a subset of the class of Emission which refers to radiation that can produce visual sensation (Light_movement:Beam);

(2.13) Medium_or_channel: a subset of the class of Physical_entities which includes physical channels, equipment, communication systems, etc., through which information is transmitted or expressed (Statement:Medium).

(3) Content: refers to cognitive or perceived content or phenomena; covers all kinds of entities, both concrete and abstract (Telling:Topic).

(4) Abstract: refers to abstract entities such as mental constructs, feelings, attitudes, among many others (Cause_to_start:Effect).

(5) Spatial_description

(5.1) Location: encompasses locations, regions or objects that mark the beginning or the destination of physical or abstract motion, the area over which a situation takes place or a participant moves, etc. (Self_motion:Area);

(5.2) Path: includes physical locations or objects which describe the trajectory that a participant follows, usually from a *Source* towards a *Destination* (Travel:Path);

(5.3) Container: a subset of the class of Physical_objects, Artifacts in particular, that are meant to be or may be used as a receptacle for something (Filling: Goal).

(6) Temporal_description

(6.1) Time: refers to a point in time when the situation occurs; includes also temporal adverbials (Being_born:Time);

(6.2) Duration: refers to time periods over which the situation occurs; includes also temporal adverbials (Taking_time:Time_length).

(7) Eventive_entities

(7.1) Eventuality: the most comprehensive class of Eventive_entities which includes activities, acts, events, processes, states, etc.; Eventualities may be performed intentionally, caused unintentionally or occurring or sustained naturally; frequently expressed by a proposition (Undergoing:Event);

(7.2.) Eventuality_activity: a subset of the class of Eventualities which includes dynamic state-of-affairs, in particular activities, acts and actions, carried out or brought about by another participant, e.g. an *Agent* or a *Cause* (Taking_time:Activity);

(7.3) Eventuality_human_act: a subset of the class of Activities which are performed specifically by persons or human entities, either intentionally or unintentionally (Purpose:Goal);

(7.4) Eventuality_process: a subset of the class of Eventualities which includes dynamic state-of-affairs, in particular processes (sustained or gradually changing phenomena), that occur on their own or are carried out or brought about by another participant, e.g. an *Agent* or a *Cause* (Cause_to_continue:Process);

(7.5) Eventuality_state: a subset of the class of Eventualities which includes non-dynamic state-of-affairs (Cause_to_end:State).

(8) Attribute

(8.1) Quantity: includes measures and quantifiable properties (Travel:Distance);

(8.2) Temperature: a subset of the class of Quantity which denotes temperature or temperature measurements (Apply_heat:Temperature);

(8.3) Currency: refers to currencies and other money related expressions (Commerse_pay:Money);

(8.4) Manner: describes how something is done or how it happens (Theft:Manner).

(9) Proposition: any content expressed by a proposition (a clause) (Attributed_information:Proposition).

Some of the restrictions, such as Abstract or Physical_entity, are defined in very general terms and often require further specification when tested on particular (subclasses of) verb synsets; other, fine-grained restrictions, such as Garment, Comestible, Sound, Currency correspond to well-defined classes of noun synsets. Temporal, quantity and other adverbials and expressions, which are marked accordingly in WordNet, are also suggested as selectional preferences where appropriate.

8. The relation *Theme* and its typology: a case study

In this section we discuss the relation *Theme* and its subtypes as an attempt at exemplifying the framework for the practical analysis and description of the relations and the selectional preferences imposed on the major participants in the conceptual structure of predicates.

Theme is a complex relation which encompasses a variety of more specific meanings. As noted in Section 6.7, it is defined between a predicate and an animate, a quasi-animate or an inanimate participant in its conceptual structure that is acted upon or affected in some way in the course of the evolving of the situation, especially by changing its location, and/or by being characterised or evaluated in terms of its state, position, condition, properties, etc. The *Theme* participant does not have control over the situation and does not undergo changes in its structure, form, function or essential properties. The relation subtypes defined in Section 6 are exemplified in Table 1 by some of their variants (corresponding to specific frame elements) and the relevant selectional preferences.

Table 1: Types and variants of the relation *Theme* with example frames and specified selectional preferences.

Relation type & their variants	Example frames	Selectional preferences
Theme		
Affected_entity	Bringing, Cause_motion, Emptying, Filling, Motion, Placing, Removing	Content: {*entity:1*}
Suspect	Arrest, Detaining, Extradition, Suspicion	Human_entity: {*person:1, individual:1...*} ∪ {*social group:1*}
Examinee	Examination	Human_entity: {*person:1, individual:1...*} ∪ {*social group:1*}
Escapee	Escaping	Volitional_animate: {*person:1, individual:1...*} ∪ {*social group:1*} ∪ {*animal:1, animate being:1…*} ∪ Vehicle: {*conveyance:3, transport:1*}
Prisoner	Imprisonment	Human: {*person:1, individual:1...*}
Evaluee	Judgment, Judgment_communication, Rewards_and_punishments	Content: {*entity:1*}
Sought_entity	Locating, Seeking, Scouring	Content: {*entity:1*}
Chosen	Choosing	Content: {*entity:1*}
Projectile	Shooting_projectiles	Projectile: {*projectile:1, missile:2*}
Clothing	Dressing, Undressing, Wearing	Garment: {*clothing:1, article of clothing:1...*}
Fluid	Cause_fluidic_motion, Fluidic_motion	Fluid: {*fluid:2*} ∪ {*body of water:1, water:2*}
Astronomical_entity	Sidereal_appearance	Astronomical_entity: {*celestial body:1, heavenly body:1*}
Theme.Situation		
Activity	Activity_start, Activity_stop, Leadership, Rising_to_a_challenge, Subversion, Intentionally_act, Justifying	Eventuality: {*event:1*} ∪ noun.act ∪ noun.event ∪ noun.process ∪ noun.phenomenon ∪ Proposition
Phenomenon	Becoming_aware, Cause_to_perceive, Evoking, Perception, Scrutinizing_for	Content: {*entity:1*} ∪ Proposition

Event	Attending, Desiring, Making_arrangements, Participation, Preventing_or_letting, Undergoing	Eventuality: *{event:1}* ∪ noun.act ∪ noun.event ∪ noun.process ∪ noun.phenomenon ∪ Proposition
Process	Process_start, Process_continue, Process_end	Eventuality: *{event:1}* ∪ noun.act ∪ noun.event ∪ noun.process ∪ noun.phenomenon ∪ Proposition
Theme.Information_content		
Message	Communication, Encoding, Questioning, Statement, Telling	Content: *{entity:1}* ∪ Proposition
Documents	Submitting_documents	Documents: *{message:1}* ∪ *{document:1, written document, papers:1}*
Opinion	Opinion, Be_in_agreement_on_assessment	Content: *{entity:1}* ∪ Proposition
Mental_content	Memorization, Remembering_information	Content: *{entity:1}* ∪ Proposition
Theme.Emitted_entity		
Sound	Make_noise, Sound_movement	Sound: *{sound:1}* ∪ *{sound:2, auditory sensation:1}* ∪ *{sound:3}* ∪ *{sound:4}* ∪ *{phone:2, speech sound:1, sound:6}*
Light	Location_of_light	Light: *{light:1, visible light:1, visible radiation:1}*
Beam	Light_movement	Light: *{light:1, visible light:1, visible radiation:1}*
Emission	Emitting, Emanating	Emission: *{radiation:1}* ∪ *{heat:1, heat energy:1}* ∪ *{fluid:2}* ∪ *{olfactory property:1, smell:2, aroma:1...}* ∪ *{secretion:2}*
Precipitation	Precipitation	Precipitation: *{precipitation:3, downfall:2}*

Excreta	Excreting	Excreta: *{body waste:1, excretion:2...}*
Theme.Transferred_possession		
Goods	Theft, Commerce_scenario, Commerce_sell, Commerce_buy, Commerce_pay	Content: {*entity:1*}
Money	Commerce_scenario, Commerce_sell, Commerce_buy, Commerce_pay,	Currency: {*currency:1*} ∪ Quantity
Possession	Possession, Losing	Content: {*entity:1*}
Theme.Body_part		
Body_part	Body_movement	Body_part: {*external body part:1*} ∪ meronyms

8.1. Theme

The general meaning of the relation *Theme* is illustrated in Example 18. In particular, the sentences exemplify a relation between a predicate and an animate (Example 18h, i) or an inanimate (Example 18a–g, j), including an abstract (Example 18j) participant that is acted upon or affected in the course of the situation so that it undergoes a change of location (Example 18a–f), position (Example 18g) or state (Example 18h), or is subject to evaluation (Example 18i, j) without changes in its structure, form, integrity, function, properties, etc.

Example 18. Example of *Theme*.

(18a) A customer **spilled [his coffee]**$_{\text{THEME.AFFECTED\_ENTITY}}$ on the floor.

(18b) He **coated** the wall **[with paint]**$_{\text{THEME.AFFECTED\_ENTITY}}$.

(18c) Jess **filled** the bottle **[with juice]**$_{\text{THEME.AFFECTED\_ENTITY}}$.

(18d) She **filled** the bag **[with work stuff]**$_{\text{THEME.AFFECTED\_ENTITY}}$.

(18e) The boy **slid [the boat]**$_{\text{THEME.AFFECTED\_ENTITY}}$ into the water.

(18f) John **gave [the beer bottle]**$_{\text{THEME.AFFECTED\_ENTITY}}$ to Maria.

(18g) John **turned [the TV]**$_{\text{THEME.AFFECTED\_ENTITY}}$ **around** towards the wall.

(18h) He **unleashed [the dogs]**$_{\text{THEME.AFFECTED\_ENTITY}}$ on the burglar.

(18i) The council **approved [us]**$_{\text{THEME.EVALUEE}}$ for the job.

(18j) The council **approved [our application]**$_{\text{THEME.EVALUEE}}$.

As the relation *Theme* is involved in a wide range of frames (see Table 1), the selectional preferences imposed on the noun member of the relation are very broadly defined as Content, which is the set union of concrete (including objects, substances, living beings, artifacts, etc.) and abstract entities. This very general definition of the preferences is further supported by the fact that in many cases a verb may select either a physical (including animate) or an abstract entity (compare Example 18i and 18j). The selectional preferences are more narrowly specified for particular classes and subclasses of verbs or even for individual verbs. Consider, for instance, the verb synset {*rubberise:1; rubber:1*} 'coat or impregnate with rubber', which enters in the *Theme* relation with a very small coherent subtree of noun synsets stemming from the synset {*rubber:2, synthetic rubber:1*} 'any of various synthetic elastic materials whose properties resemble natural rubber'.

8.2. Theme.Situation

Theme.Situation is a subtype of the relation which holds between a predicate and a participant denoting an act, activity, event, process, phenomenon, etc. that is carried out (Example 19a), acted upon (Example 19b), affected (Example 19c, d) or characterised (Example 19e) in some way in the course of the situation. The *Theme* participant may also be a situation that is perceived through the senses or the mind (Example 19f) or is the object of cognitive activity (Example 19g).

Example 19. Example of *Theme.Situation*.

(19a) He **engaged [in the president's political campaign]**$_{\text{THEME.SITUATION.EVENT}}$.

(19b) He **arranged [a meeting with the party]**$_{\text{THEME.SITUATION.EVENT}}$.

(19c) They **prevented [a catastrophe]**$_{\text{THEME.SITUATION.EVENT}}$.

(19d) He **stopped** me **[from making a mistake]**$_{\text{THEME.SITUATION.EVENT}}$.

(19e) He **justified [their behaviour]**$_{\text{THEME.SITUATION.ACT}}$.

(19f) She **anticipated [the storm]**$_{\text{THEME.SITUATION.PHENOMENON}}$.

(19g) We **hoped for [a nice day out]**$_{\text{THEME.SITUATION.EVENT}}$.

The selectional preferences for most of the variants of this subtype of the relation *Theme* are ontologically distinct from the remaining subtypes as they encompass various kinds of state-of-affairs that may be defined as Eventualities. The specification of these preferences in terms of WordNet classes is represented by the set union of the WordNet subrees denoting eventualities, particularly the tree stemming from the synset {*event:1*}, and the classes of nouns belonging to several relevant semantic primitives: noun.act ∪ noun.event ∪ noun.process ∪ noun.phenomenon. These preferences may be further narrowed down for particular verb classes (e.g. processes are distinct from acts and activities). In addition, eventualities may be expressed as propositions (syntactically – as clauses), a peculiarity which remains beyond the scope of the proposed description. In the case of Phenomenon, the entity perceived may be any physical or abstract entity and hence is additionally specified as Content.

8.3. Theme.Information_content

Theme.Information_content is a subtype of the relation which holds between a predicate and a participant denoting a piece of information – a written or a spoken message (Example 20a–c) or other types of text (Example 20f); mental content (Example 20d); opinions (Example 20e), etc. – that is conveyed or exchanged in the course of the situation.

Example 20. Example of *Theme.Information_content.*

(20a) Knowing this, she **phrased [her request]**$_{\text{THEME.INFORMATION\_CONTENT.MESSAGE}}$ in more casual terms.

(20b) She **said [that he had told her to do it]**$_{\text{THEME.INFORMATION\_CONTENT.MESSAGE}}$.

(20c) He **was not informed [of his right to speak to his attorney]**$_{\text{THEME.INFORMATION\_CONTENT.MESSAGE}}$.

(20d) If I **remember** correctly **[the price]**$_{\text{THEME.INFORMATION\_CONTENT.MENTAL\_CONTENT}}$, this must have cost you a fortune.

(20e) Fortunately, my kids **agree [in their taste in music]**$_{\text{THEME.INFORMATION\_CONTENT.OPINION}}$.

(20f) Make sure you **submit [the application form]**$_{\text{THEME.INFORMATION\_CONTENT.DOCUMENTS}}$ before the deadline.

As the information content may refer to any entity, the selectional preferences for *Theme.Information_content* is specified as Content; it may also be expressed by a proposition (Example 20b). The preferences are narrowed down for particular classes of verbs, e.g. the preference Documents (Example 20f).

8.4. Theme.Emitted_entity

Theme.Emitted_entity expresses a relation between a predicate and a participant denoting a substance or material or another physical entity or property (such as energy) that is emitted in the course of the situation (a process, a phenomenon, etc.) from a *Source*, possibly as a result of the action of a *Cause.Effector* or an *Agent*. The specific cases cover natural phenomena or otherwise produced emissions of light (Example 21a), sound (Example 21b), smell (Example 21c), precipitation (Example 21d), etc., as well as substances, such as secretions (Example 21e) or body waste.

Example 21. Example of *Theme.Emitted_entity.*

(21a) The star **shone [with a dimmed light]**$_{\text{THEME.EMITTED\_ENTITY.LIGHT}}$.

(21b) He **laughed [with a deep creepy voice]**$_{\text{THEME.EMITTED\_ENTITY.SOUND}}$.

(21c) The room **reeked [with the sweet smell of caramel]**$_{\text{THEME.EMITTED\_ENTITY.SMELL}}$.

(21d) **[Light rain]**$_{\text{THEME.EMITTED\_ENTITY.PRECIPITATION}}$ **drizzled**, slowly moistening the ground.

(21e) These animals **secrete [special substances called pheromones]**$_{\text{THEME.EMITTED\_ENTITY.EMISSION}}$.

The selectional preferences for this subtype of the relation are defined more narrowly in terms of coherent subclasses of nouns, even in the case of the most inclusive one {*radiation:1*} 'energy that is radiated or transmitted in the form of rays or waves or particles'. In the case of Sound, we specify the relevant preferences in terms of the set union of several closely related meanings of 'sound' – {*sound:1*} 'the particular auditory effect produced by a given cause' (noun.attribute), e.g. *the sound of music*; {*sound:2, auditory sensation:1*} 'the subjective sensation of hearing something' (noun.cognition), e.g. *They heard strange sounds*; {*sound:3*} 'mechanical vibrations transmitted by an elastic medium (noun.phenomenon), e.g. *The motion of the wave makes a sound*; {*sound:4*} 'the sudden occurrence of an audible event' (noun.event), e.g. *The sound startled them*; {phone:2, speech sound:1, sound:6} '(phonetics) an individual sound unit of speech without concern as to whether or not it is a phoneme of some language' (noun.communication), e.g. *The tribesmen articulated strange sounds*. This illustrates the fact that even though a semantic preference may be very coherent, it may be expressed by related noun classes that pertain to different parts of the WordNet structure, thus posing a challenge with respect to the exhaustive identification of the relevant (sub)classes or groupings of synsets that represent the preferences.

8.5. Theme.Transferred_possession

Theme.Transferred_possession is a relation defined between a predicate and a participant denoting a physical or an abstract entity, such as money (Example 22a), goods (Example 22b–e), possessions (Example 22f), that is the object of exchange or transfer of possession between participants in the situation. We should also note that part of the participants that enter into this relation may be involved in closely related situations in a different way as usually the transferred possession is given out in exchange for something else. This is the case with Money and Goods, either of which, depending on the frame, may be the *Theme* (compare Example 22a and 22b).

Example 22. Example of *Theme.Transferred_possession*.

(22a) He **paid [a lot of money]**$_{\text{THEME.TRANSFERRED\_POSSESSION.MONEY}}$ [for that sofa].

(22b) He **bought [the sofa]**$_{\text{THEME.TRANSFERRED\_POSSESSION..GOODS}}$ [for a lot of money].

(22c) They wanted to **buy [the copyright]**$_{\text{THEME.TRANSFERRED\_POSSESSION.GOODS}}$.

(22d) They tried to **buy [our loyalty]**$_{\text{THEME.TRANSFERRED\_POSSESSION.GOODS}}$.

(22e) **[New music styles]**$_{\text{THEME.TRANSFERRED\_POSSESSION.GOODS}}$ **are exported** from Britain around the world.

(22f) We **don't own [our house]**$_{\text{THEME.TRANSFERRED\_POSSESSION.POSSESSION}}$.

The selectional preferences for *Theme.Transferred_possession* are formulated very broadly (specified as Content) for a couple of reasons. First of all, as defined in the respective frames in FrameNet and as observed in the data, Goods and Possessions may be anything which is exchanged for money and include not only artifacts or physical objects or substances, but also abstract things, such as services, time, labour, intellectual property (Example 22c) or products (Example 22e). In addition, the metaphorical uses (Example 22d) represent an incoherent and inclusive class. Therefore, where possible, narrower preferences must be defined for particular groupings of verbs entering into this relation. A very specific preference, Currency, is defined for Money; at the same time the relevant participant (to which the relation points) may also be expressed by numerical expressions denoting quantities of money – which cannot be cast in terms of WordNet classes.

8.6. Theme.Body_part

Theme.Body_part is a relation between a predicate and a participant that is part of the body of an *Agent* who affects or acts upon it, especially by changing its position or spatial configuration with respect to the remaining part of the body; the body part does not undergo any change in structure, shape, function, etc. In some cases this participant must be expressed (Example 23a, b, f), while in others it may be left out as it is implicit in the meaning of the verb (Example 23c, d) or as the movement involves the whole body (compare Example 23e with 23f); with some verbs the body part may not be expressed at all (Example 23g).

Example 23. Example of *Theme.Body_part.*

(23a) They **rolled [their eyes]**$_{\text{THEME.BODY\_PART}}$ to the ceiling.

(23b) The children were **swinging [their legs]**$_{\text{THEME.BODY\_PART}}$.

(23c) She **nodded [her head]**$_{\text{THEME.BODY\_PART}}$ understandingly. — She **nodded** understandingly.

(23d) The mother **waved [her hand]**$_{\text{THEME.BODY\_PART}}$ at them. — The mother **waved** at them.

(23e) He quickly **ducked down** to avoid the blow.

(23f) He **ducked [his face]**$_{\text{THEME.BODY\_PART}}$ into her shoulder and started crying.

(23g) He **yawned**. — *He **yawned** [his mouth].

While the general selectional preference Body_part is defined in terms of the WordNet subtree stemming from {*body part:1*}, it is obvious that in the case of *Theme.Body_part* a stricter definition is required, which we define as the set union of the subtree stemming from {*external body part:1*} – including *head*, *leg*, *neck*, among others – and organs and structures that represent meronyms of external body parts, such as *eye*, *ear*, *mouth*, etc. In addition, many of the verbs involve even narrower preferences: *roll* (eyes), *nod* (head), *crane* (head, neck) and so forth.

9. Conclusions and future work

The formalisation of predicate-argument relations and selectional preferences is part of an effort directed to the enrichment of WordNet's relational structure with relations between verbs and classes of nouns corresponding to participants and other elements in their conceptual structure.

Working out a typology and a hierarchy for each of the major relations (*Agent, Patient, Experiencer*, among others), following and extending the model presented in Section 8, is essential with a view to the specification of appropriate selectional preferences.

This research may contribute both to theoretical and comparative or contrastive linguistic studies and to the implementation of methods for identification of predicate-argument structures in running text, an important NLP task with applications in semantic analysis, semantic role labelling, word sense disambiguation, language understanding and generation and machine translation.

References

Agirre, Martinez 2002: Agirre, E., D. Martinez. Integrating Selectional Preferences in WordNet. – In: *Proceedings of the First International WordNet Conference*, Mysore (India), 2002, pp. 1 – 9.

Baker 2008: Baker, C. FrameNet, Present and Future. – In: *Proceedings of First International Conference on Global Interoperability for Language Resources (ICGL-2008)*, Hong Kong, January 2008, pp. 12 – 17.

Baker et al. 1998: Baker, C. F., C. J. Fillmore, J. B. Lowe. The Berkeley FrameNet Project. – In: *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (COLING-ACL'98)*, 10 – 14 August 1998, Montreal, Canada, pp. 86 – 90.

Baker, Fellbaum 2009: Baker, C. F., C. Fellbaum. WordNet and FrameNet as Complementary Resources for Annotation. – In: *Proceedings of the Third Linguistic Annotation Workshop (ACL-IJCNLP'09)*, ACL, Stroudsburg, PA, USA, pp. 125 – 129.

Chafe 1970: Chafe, W. L. (1970). *Meaning and the Structure of Language*, University Press, Chicago.

Di Fabio et al. 2019: Di Fabio, A., S. Conia, R. Navigli. VerbAtlas: a Novel Large-Scale Verbal Semantic Resource and its Application to Semantic Role Labeling. – In: *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, 2019, Association for Computational Linguistics, pp. 627 – 637.

Fellbaum 1998: Fellbaum, C. (Ed.) *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

Fellbaum, Baker 2008: Fellbaum, C., C. F. Baker. Can WordNet and FrameNet be made "interoperable"? – In: Jonathan Webster, Nancy Ide, and Alex Chengyu Fang (Eds.) *Proceedings of The First International Conference on Global Interoperability for Language Resources*, pp. 67 – 74, Hong Kong, City University.

Fellbaum et al. 2009: Fellbaum, C., A. Osherson, P. E. Clark. Putting Semantics into WordNet's "Morphosemantic" Links. – In: *Proceedings of the Third Language and Technology Conference*, Poznan, Poland. [Reprinted in: *Responding to Information Society*

Challenges: New Advances in Human Language Technologies. Springer Lecture Notes in Informatics], vol. 5603, pp. 350 – 358.
Ferrandez et al. 2010: Ferrandez, O., M. Ellsworth, R. Munoz, C. F. Baker. Aligning FrameNet and WordNet Based on Semantic Neighborhoods. – In: *Proceedings of the 7th Conference on Language Resources and Evaluation (LREC 2010)*, 17 – 23 May 2010, Valletta, Malta, pp. 310 – 314.
Fillmore 1982: Fillmore, C. J. Frame Semantics. – In: The Linguistic Society of Korea (Eds.) *Linguistics in the Morning Calm*. Seoul: Hanshin, pp. 111 – 137.
Foley, Van Valin 1984: Foley, W., R. Van Valin. *Functional Syntax and Universal grammar*, Cambridge University Press.
Goldberg 1994: Goldberg, A. *Constructions: A Construction Grammar Approach to Argument Structure*, University of Chicago Press.
Kipper et al. 2008: Kipper, K., A. Korhonen, N. Ryant, M. Palmer. A Large-scale Classification of English Verbs. – In: *Language Resources and Evaluation*, vol. 42 (1), pp. 21 – 40.
Kipper-Schuler 2005: Kipper-Schuler, K. 2005. *VerbNet: A Broad-coverage, Comprehensive Verb Lexicon*. PhD. Thesis. Computer and Information Science Dept., University of Pennsylvania. Philadelphia.
Koeva 2008: Koeva, S. Derivational and Morphosemantic Relations in Bulgarian Wordnet. – In: *Intelligent Information Systems*, XVI, Warsaw, Academic Publishing House, pp. 359 – 389.
Koeva 2010: Коева, С. *Българският ФреймНет 2010*. София, 2010. [Koeva, S. Balgarskiyat FreymNet 2010. Sofia, 2010.]
Koeva et al. 2016: Koeva, S., S. Leseva, I. Stoyanova, T. Dimitrova, M. Todorova. Automatic Prediction of Morphosemantic Relations. – In: *Proceedings of the Eighth Global WordNet Conference*, Research Institute for Artificial Intelligence – Romanian Academy, 2016, pp. 168 – 176.
Koeva et al. 2020: Koeva, S., T. Dimitrova, V. Stefanova, D. Hristov. Towards Conceptual Frames. (this volume)
Lakova 2015: Лакова, М. Извънизреченските семантични роли в съвременния български книжовен език. – В: *Съпоставително езикознание,* 1, 2015, с. 35 – 59. [Lakova, M. Izvanizrechenskite semantichni roli v savremenniya balgarski knizhoven ezik. – In: *Sapostavitelno ezikoznanie*, 1, 2015, pp. 35 – 59.]
Laparra, Rigau 2010: Laparra, E., G. Rigau. eXtended WordFrameNet. – In: *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010)*, 17 – 3 May 2010, Valletta, Malta, pp. 1214 – 1219.
Leseva et al. 2018: Лесева, Св., Ив. Стоянова, Хр. Кукова, М. Тодорова. Интегриране на субкатегоризационна информация в релационната структура на УърдНет. – В: *Български език*, 65, 2, 2018, с. 11 – 40. [Leseva, S., I. Stoyanova, H. Kukova, M. Todorova. Integrirane na subkategorizatsionna informatsiya v relatsionnata struktura na Uardnet. – In: *Balgarski ezik*, 65, 2, 2018, pp. 11 – 40.]
Leseva et al. 2019: Leseva, S, I. Stoyanova, M. Todorova, H. Kukova. A Theoretical Overview of Conceptual Frames and Semantic Restrictions on Frame Elements. – In: *Linguistique Balkaniqu*e, LVIII, 2, 2019, pp. 172 – 186.
Leseva et al. 2020: Leseva, S, I. Stoyanova, M. Todorova, H. Kukova. A Semantic Description of the Combinability between Verbs and Nouns (on Material from Bulgarian and English). – In: *Чуждоезиково обучение*, 47, 2, 2020, с. 115 – 128. [*Chuzhdoezikovo obuchenie*, 47, 2, 2020, pp. 115 – 128.]
Leseva, Stoyanova 2019: Leseva, S., I. Stoyanova. Enhancing Conceptual Description through Resource Linking and Exploration of Semantic Relations. – In: *Proceedings of the Tenth*

Global WordNet Conference, 23 – 27 July 2019, Wrocław, Poland. Oficyna Wydawnicza Politechniki Wrocławskiej, 2019, pp. 280 – 289.
Leseva, Stoyanova 2020: Leseva, S., I. Stoyanova. Beyond Lexical and Semantic Resources: Linking WordNet with FrameNet and Enhancing Synsets with Conceptual Frames. (this volume)
Levin 1993: Levin, B. *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press.
Longacre 1976: Longacre, R. E. *An Anatomy of Speech Notions*, Peter de Ridder Press.
Miller 1995: Miller, G. A. WordNet: A Lexical Database for English. – In: *Communications of the ACM*, 38 (11), pp. 39 – 41.
Palmer 2009: Palmer, M. Semlink: Linking PropBank, VerbNet and FrameNet. – In: *Proceedings of the Generative Lexicon Conference*, pp. 9 – 15.
Palmer et al. 2014: Palmer, M., C. Bonial, D. McCarthy. SemLink+: FrameNet, VerbNet and Event Ontologies. – In: *Proceedings of Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929–2014)*, Baltimore, USA, ACL, pp. 13 – 17.
Pinker 1989: Pinker, S. *Learnability and Cognition: The Acquisition of Argument Structure*, MIT Press.
Ruppenhofer et al. 2016: Ruppenhofer, J., M. Ellsworth, M. R. L. Petruck, C. R. Johnson, C. F. Baker, J. Scheffczyk. *FrameNet II: Extended Theory and Practice* (Revised November 1, 2016). <https://framenet2.icsi.berkeley.edu/docs/r1.7/book.pdf> [01.06.2020]
Shi, Mihalcea 2005: Shi, L., R. Mihalcea. Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robust Semantic Parsing. – In: Gelbukh, A. (Ed.). *Computational Linguistics and Intelligent Text Processing (CICLing 2005)*. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, vol. 3406, pp. 100 – 111.
Tonelli, Pighin 2009: Tonelli, S., D. Pighin. New Features for FrameNet – WordNet Mapping. – In: *Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL'09)*, Boulder, USA.
Urešová et al. 2020: Urešová, Z., E. Fucíková, E. Hajicová, J. Hajic. SynSemClass Linked Lexicon: Mapping Synonymy between Languages. – In: *Proceedings of the Globalex Workshop on Linked Lexicography*, Language Resources and Evaluation Conference (LREC 2020), Marseille, 11–16 May 2020, pp. 10 – 19.
Van Valin, LaPolla 1997: Van Valin Jr., R. D., R. J. LaPolla. *Syntax: Structure, Meaning and Function*, Cambridge University Press.

Svetlozara Leseva
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria
zarka@dcl.bas.bg

Ivelina Stoyanova
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria
iva@dcl.bas.bg

Maria Todorova
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria
maria@dcl.bas.bg

Hristina Kukova
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria
hristina@dcl.bas.bg

TOWARDS CONCEPTUAL FRAMES

Svetla Koeva, Tsvetana Dimitrova,
Valentina Stefanova, Dimitar Hristov

Institute for Bulgarian Language *Prof. Lyubomir Andreychin*
Bulgarian Academy of Sciences

Abstract: This study presents an effort in enriching the Princeton WordNet with information from the verb patterns (predicate-argument structures) from another resource – the Pattern Dictionary of English Verbs (PDEV) – and the semantic types from the Corpus Pattern Ontology (CPA) which describe the arguments of these patterns. The PDEV verb patterns were automatically mapped to the WordNet sentence frames thus adding information about the character of the arguments. The resulting patterns are conceptual frames whose arguments were specified for a set of lexical units – the semantic types assigned to WordNet noun synsets.

Keywords: *lexical-semantic resources, WordNet, frame semantics, syntax*

1. Introduction

We present a series of papers – in chapters – describing a concerted effort on enriching the WordNet with information that has been encoded into the verb patterns of the Pattern Dictionary of English Verbs (PDEV) – predicate-argument structures with semantic values for the arguments that are selected from a set of semantic types extracted on the basis of corpus data and hierarchically organised into the Corpus Pattern Analysis (CPA) ontology.

The first chapter deals with the enrichment of WordNet data through merging of WordNet concepts and CPA semantic types. The 253 CPA semantic types were mapped to the respective WordNet concepts, and as a result the hyponyms of a synset to which a CPA semantic type is mapped, inherited not only the respective WordNet semantic primitive but also the CPA semantic type.

The second chapter describes the mapping of the WordNet and the PDEV, with 2,593 sentence frames in WordNet being automatically extended with information from 2,904 unique PDEV verb patterns. The extended sentence frames assigned to top hypernyms were manually validated.

The third chapter reports on the process of manual validation of verb patterns that have been automatically assigned to the WordNet verb synsets. 4,084 patterns were approved, 1,568 new patterns were manually assigned and 2,815 inappropriate patterns were removed. The chapter introduces the notion of conceptual frame: a FrameNet semantic frame, whose core frame elements are specified for a set of lexical units. The granular semantic types assigned to WordNet noun synsets defined the sets of WordNet literals (lexical units) appropriate to express the core frame elements.

The results of our work are publicly available and can be accessed at: http://dcl.bas.bg/semantichni-mrezhi/.

2. Mapping WordNet Concepts with CPA Ontology[30]

2.1. Introduction to the mapping of WordNet concepts with CPA ontology

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

WordNet synsets are classified into semantic primitives (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 test and enrich semantic relations between noun and verb synsets (of the type of morphosemantic relations – Agent, Undergoer, Instrument, Event, etc. – that link verb-noun pairs of synsets that contain derivationally related literals) (Fellbaum et al. 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 can collocate 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 chapter, 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 organised within the so-called CPA ontology.

Our goal is 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 chapter is organised as follows: in section 2.2, we present our motivation for the work before discussing different attempts at semantic classification of nouns in

[30] This chapter was first published as: Koeva, S., T. Dimitrova, V. Stefanova, D. Hristov. Mapping WordNet concepts with CPA ontology. – In: *Proceedings of the 9th Global WordNet Conference (GWC'2018)*, Global WordNet Association, Singapore, 2018, pp. 70 – 77.

section 2.3. Section 2.4 briefly presents the CPA ontology, while section 2.5 outlines some issues with the WordNet noun hierarchy. The effort at mapping CPA semantic types and WordNet concepts is discussed in section 2.6, with a comparison between the two structures in 2.7 and some conclusions.

2.2. 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 (1b) and (1c) may be specified as animate.

(1)

a. {purr:1[31], 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 organisation 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 hypothesise 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 (further specifying 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.

---

[31] The numbers of the literals – where available throughout this paper – follow those applied in the database used by the viewer Hydra, at: http://dcl.bas.bg/bulnet/.

Previous work on mixing resources and enriching the information on semantic and syntactic behaviour of verbs encoded in WordNet builds upon resources – one or more than one – that use Levin'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, 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.

2.3. 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, 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's distributional hypothesis (1954) 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. This 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 generalisation than the type. These are namely: mass and countable; plus three additional for subcategorisation: trans (nouns with thematic complements introduced by the preposition *de*); intrans (nouns without complements); pcomp (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.

2.4. 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 organised 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.

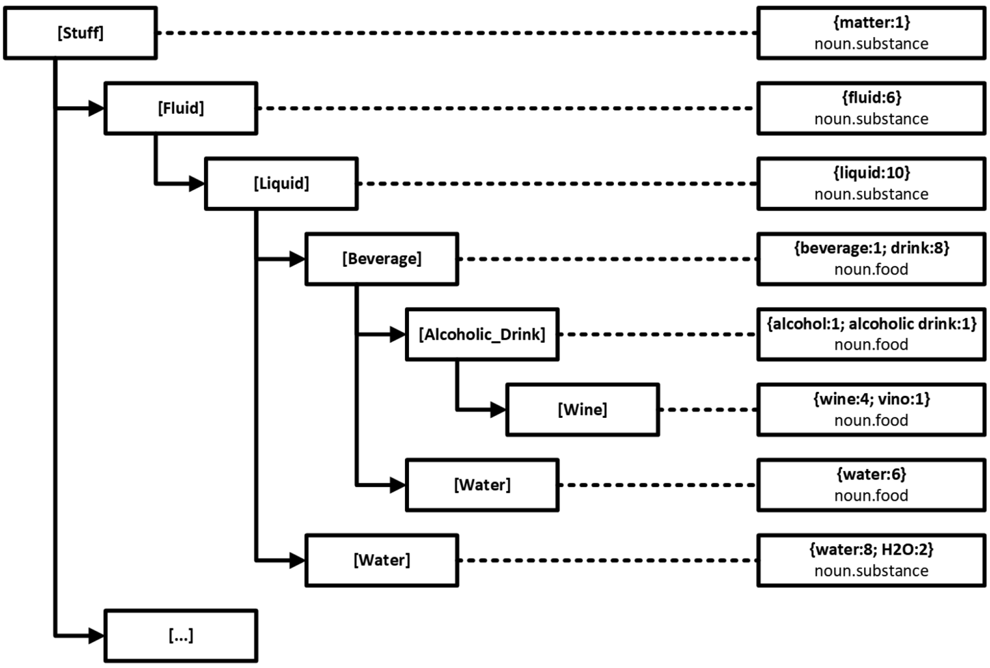

Figure 1: Part of the CPA ontology

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, 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, 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 can be 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 ontology of semantic types.

2.5. WordNet noun hierarchy

Noun synsets in WordNet are organised 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, 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 hypernym, 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, there is the EuroWordNet top ontology which contains 63 semantic primitives (Vossen 2002). The ontology is designed to help the encoding of WordNet semantic relations in a uniform way. The 1st Order Entities are distinguished in terms of main ways of conceptualising or classifying a concrete entity (Pustejovsky 1995):

Origin, Form, Composition and Function. Further, Origin is divided into Natural and Artifact, and Natural – in 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.

2.6. 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 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 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 are 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.

For example, all hyponyms of the WordNet synset {location:1} ‘a point or extent in space’ are classified 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 better prediction for the words connectivity and thus 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 primitive 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 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 online[32], marked with the XML tag CPA in the WordNet noun synsets.

2.7. 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

[32] http://dcl.bas.bg/PWN_CPA/ [22 May 2020]

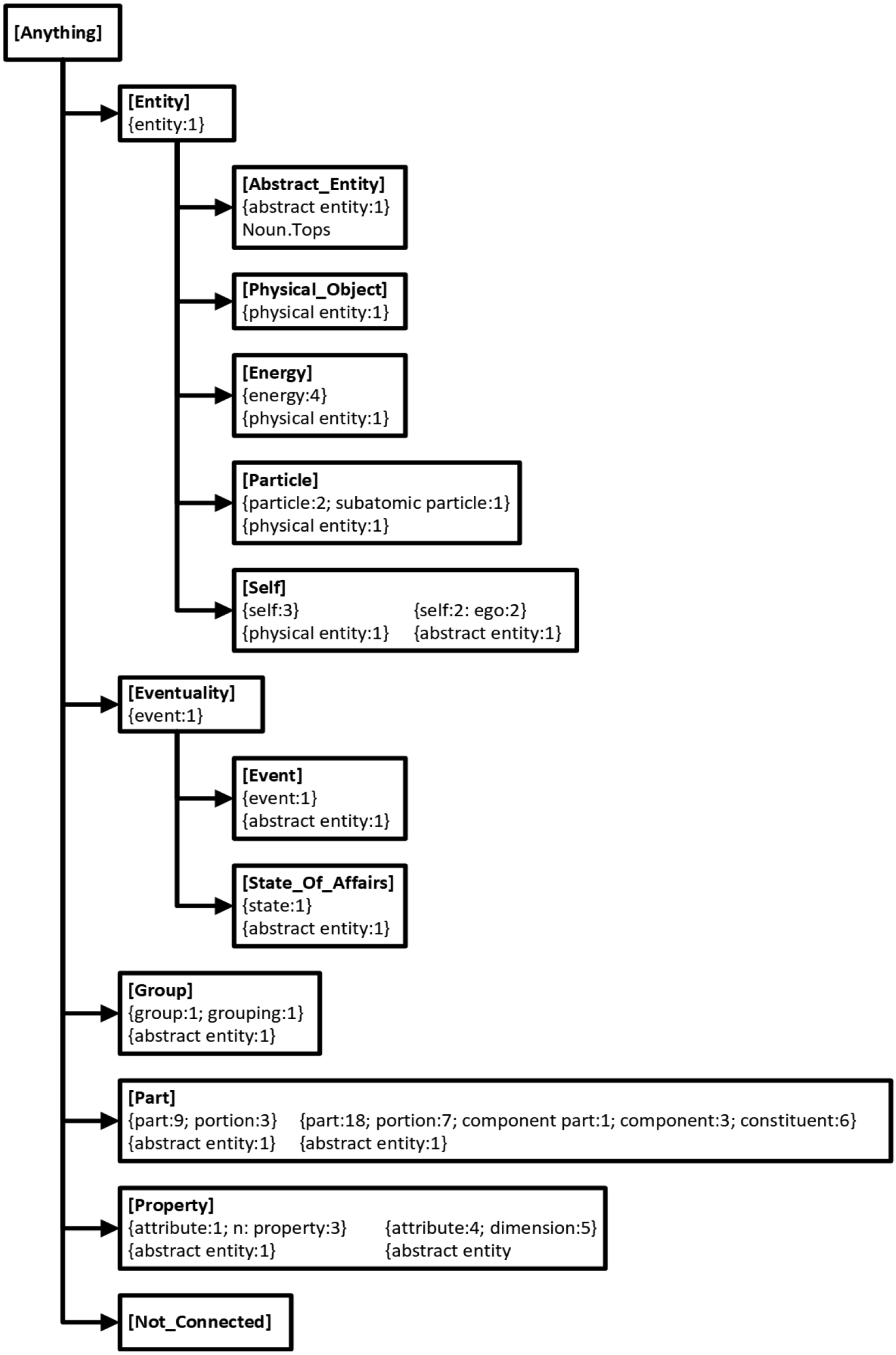

Figure 2: Matching

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 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 sub-hierarchies with probable noun synset(s), which are linked to either of the two noun.Tops; some types 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.

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 hypernymy paths:

[Artwork]
{artwork:1; art:4; graphics:2; nontextual matter:1} ← {visual communication:1} ← {n: communication:1} ← {abstraction:1; abstract entity:1}
{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} and {substance:2}

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

[Musical Instrument]
{musical instrument:1; instrument:6} ← {device:2} ← {instrumentality:1; instrumentation: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 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 drown:

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).

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 expressed within the WordNet semantic network became explicit.

3. Enriching WordNet with Frame Semantics[33]

3.1. Introduction to the enriching of WordNet with frame semantics

This chapter reports the mapping of two lexical-semantic resources – WordNet (Miller et al. 1990; Miller 1995; Fellbaum 1998) and the Pattern Dictionary of English Verbs (PDEV) (Hanks 2004; Hanks, Pustejovsky 2005; Hanks 2008) with the aim of enriching WordNet. Although the two resources are designed to collect and model semantic information, the information is organised differently.

PDEV is built on the basis of the lexicocentric Theory of Norms and Exploitations (Hanks 2013) and uses a mechanism called Corpus Pattern Analysis (CPA) for mapping meanings onto words in a corpus (the British National Corpus). PDEV

[33] This chapter was first published as: Koeva, S., D. Hristov, T. Dimitrova, V. Stefanova. Enriching WordNet with Frame Semantics. – In: *Доклади от Международната годишна конференция на Института за български език „Проф. Любомир Андрейчин“ (София, 14 – 15 май 2019 година)*. София: Издателство на БАН „Проф. Марин Дринов“, 2019, с. 300 – 308. *[Dokladi ot Mezhdunarodnata godishna konferentsiya na Instituta za balgarski ezik „Prof. Lyubomir Andreychin“ (Sofiya, 14 – 15 may 2019 godina)*. Sofia: Izdatelstvo na BAN „Prof. Marin Drinov“, 2019, pp. 300 – 308]

consists of verb patterns in the form of predicate-argument structures, with arguments being defined in terms of their semantic type. These semantic types are organised within a hierarchical CPA ontology and both verb patterns and semantic types are defined on the basis of patterns and usages found in specific contexts in a corpus.

WordNet, on the other hand, contains structured lexical-semantic information based on cognitive concepts. WordNet verb synsets further contain the so-called verb (sentence) frames that describe the set of (obligatory) arguments a verb may take to form a simple sentence, as well as the reference value (human vs. non-human) and form (*-ing*, prepositional phrase (PP), *to*-infinitive, *whether*-clause, *that*-clause, etc.) of the arguments (e.g., *Something ----s; Somebody ----s; Somebody ----s Something*; *Somebody ----s to INFINITIVE*, etc.). The frames are associated with a synset (a verb synset must have at least one frame) if they refer to all literals in the synset, otherwise frames are attributed only to literals. The frames, however, are not detailed enough (they do not impose semantic restrictions on the arguments and in some cases do not contain all relevant arguments). The syntactic information in WordNet is scarce because it was "conceived as a semantic database only", as pointed out by Fellbaum (1998: 11).

In our effort, we aim to make explicit existing relations between WordNet concepts, e.g., *[Person]* ***graduates*** *from [University or School]* – such relations are retrievable from the network, but there are no direct links yet. We lean on the fact that the semantics of the top hypernym subsumes the semantics of its hyponyms. Thus, we focus on enriching the WordNet top hypernyms' verb sentence frames with semantic information so as to specify their semantic compatibility (semantic types of arguments). To this end, we automatically mapped the PDEV verb patterns to the WordNet verb synsets and cross-validated the PDEV patterns and the WordNet verb frames to provide explicit knowledge in WordNet for the semantic compatibility of verbs with sets of entities (noun synsets) belonging to particular semantic types.

3.2. Related work

A couple of previous efforts aimed at linking lexical-semantic resources, simply combine the information available in each of them (resources with Levin's classes, VerbNet, FrameNet, PropBank).

Korhonen (2002) proposes a semi-automatic semantic classification of verbs in WordNet using Levin's classes (Levin 1993), Longman Dictionary of Contemporary English and Dorr's (1997) source of the Longman Dictionary of Contemporary English grammatical codes for Levin's classes. WordNet senses are classified according to Levin's classes; individual literals are classified on the basis of their WordNet senses and hyponyms are classified under the semantic class of the hypernym.

Green et al. (2001) describe an effort in automatic mapping of entries in a database of over 4,000 English verbs (classified into Levin's classes and extended through the splitting of some classes into subclasses to form new classes) to WordNet senses. Each new class was associated with a thematic grid in the form of a predicate-argument structure. Verb entries were manually tagged with WordNet senses as training data. Probabilistic associations were made between grids and WordNet verb frames and SemCor[34] frequency, with disambiguation based on semantic similarity.

Integration of WordNet, VerbNet and FrameNet was proposed by Shi and Mihalcea (Shi, Mihalcea 2005). A unified knowledge base was built for a robust rule-based semantic parser intended to identify the semantic structures in a text. A combination of manual and automatic procedures was used to map the semantic constituents in the three resources (frames in FrameNet, semantic roles in VerbNet and semantic classes in WordNet). In the new resource, the coverage of FrameNet was extended, frame semantics was added to the VerbNet lexicon, selectional restrictions were implemented using WordNet classes.

An effort at combining sense relational knowledge available in WordNet with frames from VerbNet and corpus knowledge from PropBank was performed by Pazienza et al. (2006) to the end of creating a large set of linguistic examples of verb pairs that have a semantic relation and specific predicate-argument structures.

In some cases, the WordNet mapping with FrameNet was used for FrameNet lexical units induction in different languages (De Cao et al. 2008; Crespo, Buitelaar 2008; Tonelli, Pighin 2009).

The aim of the effort presented in this chapter is to use the alignment between WordNet synsets and PDEV patterns for cross-validation of the information encoded in the two resources and for further enrichment of WordNet with semantic relations.

3.3. PDEV and WordNet

Mapping the PDEV verb patterns and WordNet sentence frames is used for expanding WordNet semantic relations provided that: 1) The semantic types from the CPA ontology are featured as arguments of a given predicate in the PDEV patterns; 2) The WordNet noun synset hierarchy is already mapped onto the semantic type hierarchy in the CPA ontology previously undertaken to expand the WordNet semantic primitives as described by Koeva et al. (2018) (see chapter 1).

In this section, we briefly present the organisation of relevant information in the two resources.

3.3.1. PDEV verb patterns

The verb patterns in PDEV represent predicate-argument structures with semantic values for the arguments (selected from a set of semantic types hierarchically organised into the CPA ontology). Thus, verb senses are defined in relation to nouns that co-occur with each verb (denoting the particular verb sense) in different roles: subject, direct object, propositional object or adverbial. Different senses for a word are represented implicitly by assigning different patterns, similarly to FrameNet where different senses are represented by assigning different frames. Verb patterns

[34] https://www.sketchengine.eu/semcor-annotated-corpus/ [22 May 2020]

are organised as ordered predicate-argument structures where arguments are written in capital letters in square brackets and are assigned a semantic type.

The patterns, however, include also non-obligatory elements – written in parentheses for optionality, with alternatives in some positions – marked by a disjunctive bar, and sub-valency items, such as modifiers, e.g., [Human] | [Animate] *leads* !MOD! life | existence; determiners: [Anything] *impresses* [Human] | [Institution] (with [REFLDET] Property]), etc. Semantic types can be complemented by semantic (or contextual) roles implied by '=', as in [Human = Police Officer] – a contextual role encodes a property assigned to the noun in a specific context in the corpus (Cinkova, Hanks 2010). An example for both lexical sets (in curly brackets) and contextual roles is the following verb pattern, e.g., [Human] | [Institution] *denies* {responsibility | liability} for [Event = Bad] | knowledge of [Event = Bad] | knowledge of [Event = Bad]. A verb pattern may include semantic types from different levels of the CPA ontology but also from more general types at the top – for example, in: [Human] *attacks* [Activity] | [System] | [Anything] – the type [Anything] subsumes the others.

The PDEV verb patterns also define the syntactic realisation of some arguments such as: PPs with specific prepositions, e.g., [Human] *abstains* (from [Activity]; an -ING form, e.g., [Human] | [Institution] leads the way (in [-ING]); quotations, *that*-clauses, *wh*-clauses, e.g., [Human] *notes* #quote# | #*that*-clause# | #*wh*-clause#; etc.

The classification according to the ontology of semantic types is one of the main advantages of this resource, while the limited coverage is one of its main problems.

3.3.2. WordNet frames

In WordNet, each verb synset contains sentence frames (at least one) that describe the set of arguments the verb or verbs in the synset may take. The whole list includes[35] frames which define the number and type of arguments, respecting a general semantic distinction between HUMAN (*Somebody*) and NON-HUMAN (*Something*), and syntactic constraints on the realisation of arguments (*-ing*, *that*-CLAUSE, *to*-INFINITIVE, prepositions and their arguments (*to Somebody, from Somebody*, etc.).

Verb frames do not explicitly describe all the structures subsumed under one frame (e.g., a *that*-clause may presuppose alternation with a *wh*-clause) and no further semantic distinctions are made (for example, animals are classified as *Something*, i.e., NON-HUMAN).

Arguments may be realised by a prepositional phrase (PP), as well as by a *that*-CLAUSE or *-ing* form. A variation in the tree is observed with the addition of a PP, such as: *with Something*, *to Somebody*, *from Someone*, *on Somebody*, *on Something*, etc.

To sum up, WordNet provides intensive lexical coverage and dense semantic links, but lacks information about frame semantics and its reflection via explicit se-

[35] [REFLDET] – reflexive possessive determiner.

mantic relations. Sentence frames refer mainly to the syntactic realisation. An advantage of the resource is that the difference in the syntactic realisation is encoded on the level of the literals.

3.3.3. Verb frames within hypernym/hyponym trees

Considering the hypernym/hyponym trees, the hyponyms' sentence frames usually involve the verb frame(s) of the hypernym. This is due to the fact that although WordNet sentence frames are in general syntactic frames, they correspond to (non-explicitly encoded) semantic frames. We assume that within a hypernym/hyponym tree the semantic frames of hyponyms have to be inherited from the hypernym, thus although some arguments may remain implicit at the syntactic level, they should be present at the semantic level. Since WordNet sentence frames are in general syntactic frames, they represent differences in the syntactic realisation of the arguments of hypernyms, hyponyms and synonyms, if any.

We have observed a handful of tendencies which depend mainly on the characteristics of the subject – whether it is classified as *Somebody* or as *Something*:

If the verb frame of the top verb hypernym is *Somebody ----s*, the variation within the verb frames of the hyponym verbs is lower and the subject is animate: *Somebody ----s*, and/or *Something ----s*, where the subject is predominantly of the animal or group class (classified with the semantic primitives noun.animal or noun.group).

The subject is non-animate only when the top verb frames set include *Something ----s*.

If the top verb frame is only *Something ----s*, the appearance of *Somebody ----s* frame with a hyponym verb may be considered an error.

3.3.4. WordNet semantic primitives

Another source of semantic information about the arguments are semantic primitives (or semantic classes, cf. Miller et al. 1990). The semantic classification of WordNet nouns and verbs is consistent, but the sentence frames lack the information for further constraining the set of concepts that would saturate nominal arguments. Thus, the noun semantic primitives might be further specified if we want to establish links between a given verb and sets of nouns collocating with it (e.g., with ***read*** and its frame *Somebody ----s something*, a non-human noun can be *Something*).

We have mapped the WordNet noun synset hierarchy onto the semantic types in the CPA ontology by matching the CPA semantic types with WordNet synsets and choosing the most probable ones, with nonexhaustive results (Koeva et al. 2018 – see chapter 1). The 253 CPA semantic types are mapped to the respective WordNet concept (synset), with 199 semantic types mapped directly to one concept (e.g., [Permission] – {permission} 'approval to do something', noun.communication) and 54 semantic types mapped to more WordNet concepts (e.g., [Language] to {language} 'a systematic means of communicating by the use of sounds or conventional symbols', noun.communication; and {language} ‘the mental faculty or power of vocal communication’, noun.cognition).

WordNet semantic primitives are always preserved, and new (complementary) semantic types borrowed from the CPA ontology were added.

3.4. Mapping WordNet and PDEV

The primary attention is, therefore, focused on the mapping of the WordNet verb synsets with the PDEV patterns. We used two different methods for obtaining PDEV data – applying the provided PDEV-Lemon resource (El Maarouf et al. 2014) which employs the Lemon lexicon model (McCrae et al. 2012) and extracting the data from the PDEV website interface (web scraping).

3.4.1. PDEV-lemon and web data extraction

The PDEV-lemon consists of several large Lemon lexicon files which contain information about all the entities in the resource[36]. These entities are connected through references; e.g., the entry for a pattern references a subject argument entry, which references a semantic type (class) entry, which contains the semantic type’s name. This structure provides for a robust building and understanding of the PDEV patterns. However, a number of references are linked to non-existent data entries, leaving “holes” in the structure and resulting in incorrect patterns.

Unlike PDEV-lemon, data gathered through web data extraction is not relational and, hence, does not provide for a robust and ambiguity-free pattern building. However, the data was still separated into major pattern subparts – subject, verb form, indirect object, object, complement and adverbials, together with markers for phrasal verbs and idioms and pattern frequency; this data was enough for the purpose of the experiment. The data was extracted using PhantomJS and a Bash script for crawling the resulting HTML[37]. Resulting pattern entries were filtered and any empty patterns, patterns for phrasal verbs or idioms and any patterns with frequency under 0.1% were removed.

3.4.2. Translation of PDEV patterns to WordNet frames

A set of translation rules was applied to convert PDEV patterns into WordNet frames and to preserve information of optional pattern arguments and alternatives. The translation observes the following rules:

1. Every argument contains one or more alternatives given for the whole argument or parts of it, separated by an OR operator.
2. Verb forms are converted to “----s”, “is ----ing” and “is ----ed” (for passive voice, not present in WordNet).
3. If any of the alternatives for an argument is given as [Human], the argument is translated to “Somebody” regardless of other alternatives.

[36] http://pdev.org.uk/ PDEVLEMON.html [21 November 2018]
[37] https://pdev.org.uk [21 November 2018]

4. The PDEV argument [Body_Part] is translated into “Somebody’s (body part)” when in the position of a subject, and into “Something” otherwise.
5. The PDEV arguments [N], [NP] and [ADJ] are translated into WordNet arguments, as follows: [N] and [NP] to “Noun”; [ADJ] to “Adjective”; [N | ADJ] and [NP | ADJ] to “Adjective/Noun”.
6. The PDEV arguments “-ing” and “-ING” are translated into “VERB-ing” when in the position of an object.
7. The PDEV arguments “quote”, “to-infinitive”, “wh-clause” and “that-clause” are translated into WordNet arguments, as follows: “quote” to the alternatives “Something” and “whether INFINITIVE”; “to-infinitive” to the alternatives “INFINITIVE” and “to INFINITIVE”; “wh-clause” to the alternative “that CLAUSE” and “whether INFINITIVE”; “that-clause” to “that CLAUSE”.
8. Any other class or lexical entity is translated into “Something”.
9. For prepositional phrases, the PDEV arguments were translated to alternative arguments (as the operator OR may be used for alternatives of the preposition or the argument):
   a. Prepositions “from”, “of”, “on”, “to” and “with” are kept as such and their argument is translated according to previous rules.
   b. “into [-ING]” is translated to “into V-ing”.
   c. Any other prepositional phrase is translated to “PP”.
10. Sub-valency items are omitted.
11. If an argument is marked as optional in the PDEV pattern (by putting it in [square] brachets), an empty alternative is added to the translation of the argument.

Arguments with alternatives were then expanded. As a result, a single PDEV pattern is translated into one or more WordNet frames.

Example:

[Anything] | [Body_Part] <verbs> ([ADJ]) Gives:

Something | Somebody’s (body part) ----s Adjective | <empty> Expanded to:

Something ----s

Something ----s Adjective

Somebody’s (body part) ----s

Somebody’s (body part) ----s Adjective

3.4.3. Assigning PDEV patterns to WordNet synsets

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 ∈ S, PDEV pattern P ∈ patterns(L) can be assigned to S if and only if frames(S) ∩ translations(P) ≠ Ø.

Here patterns(L) is the set of PDEV patterns with the base verb L, frames(S) is the set of WordNet frames for the synset and translations(P) is the set of translation variants (which are WordNet frames) of the PDEV pattern P.

3.4.4. Results

As a result, the sentence frames of verb synsets in WordNet are extended with the information from the PDEV patterns and supplied with the CPA semantic types. 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 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 (that do not have hypernyms). This resulted in 453 synset-pattern assignments.

These numbers are a result of the size of the PDEV resource, which contains patterns for 1,375 English verbs; hence, not all literals in WordNet synsets have been assigned patterns from PDEV. Also, a number of PDEV patterns do not have translations that appear as frames in WordNet (which is the case with passive voice patterns).

As a result, the sentence frames of verb synsets in WordNet are extended with the information from the PDEV patterns and supplied with the CPA semantic types.

3.5. Cross-validation of results

The automatic mapping was subjected to manual validation where the following cases were observed.

3.5.1. Exact match

The exact matches were few and covered both one-place predicates, as in (1a) below, and two-place predicates, as in (1b) below.

(1)

a. {halt:5} ‘come to a halt, stop moving’
Something ----s; Somebody ----s
[Human] | [Vehicle] **halts**

b. {work:22} ‘be employed’
Somebody ----s; Somebody ----s PP
[Human] **works;**
[Human 1] **works** for [Human 2] | [Institution];
[Human] | [Institution] **works** for [State_of_Affairs]

In these cases, CPA semantic types were applied to the WordNet verb frames, allowing for the establishment of semantic relations between the verb synsets in the hypernym/hyponym tree and the noun synset to which the respective semantic type was mapped, as well as its hyponyms (e.g., (1a): *Something ----s; Somebody ----s > Vehicle ----s; Human ----s*; (1b): *Somebody ----s; Somebody ----s PP > Human ----s; Human ----s for Human | Institution; Human | Institution ----s for State_of_Affairs).*

3.5.2. Subsumption of WordNet verb frames into PDEV patterns

In most cases, WordNet verb 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 frames to match the PDEV patterns only partially.

In (2a) below, the WordNet verb frames for {appear: 2} subsume the patterns in PDEV – they do not feature time or place adverbials or PP arguments as adverbials, and PPs are optional in the pattern.

In (2b) below, the two PDEV patterns feature a PP argument (non-obligatory) while a PP argument is found with one of the two WordNet verb frames.

(2)

a. {appear:2} 'come into being or existence, or appear on the scene'
Something ----s, Somebody ----s
[Physical_Object] **appears**;
[Abstract_Entity] [State_of_Affairs] **appears**;
[Stuff] | [Physical_Object] **appears** *Time Period*;
[Human] | [Animal] **appears** (*Location*);
[Document 1] | [Image] **appears** (in [Document 2])

b. {agree:3} 'achieve harmony of opinion, feeling, or purpose'
Somebody ----s; Somebody ----s PP
[Human 1] | [Institution 1] **agrees** (with [Human 2] | with [Institution 2]) (about [Topic]);
[Human] | [Institution] **agrees** (with [Proposition] | with [Activity])

In cases where both the WordNet verb frame and the PDEV pattern were correct, but the PDEV pattern contained more syntactic information, we took the syntactic and semantic information from the PDEV pattern and the additional CPA semantic types were applied to the WordNet verb frames, e.g., [Human] *chooses* [Anything]; [Human] *chooses #wh*-clause#; [Human] *chooses #that*-clause# & 'choose' *Somebody ----s something*; *Somebody ----s that*-CLAUSE > *Human ----s Anything*; *Human ----s that*-CLAUSE; *Human ----s wh*-CLAUSE.

3.5.3. More difficult cases

Another group featured cases where only one of the WordNet frames matched a PDEV pattern, as in (3a) below, or matched more than one argument in the WordNet frames: in (3b) below, [Anything] subsumes [Something] and [Somebody].

(3)

a. {shame:4} 'bring shame or dishonor upon'
Somebody ----s something; Somebody ----s somebody; Something ----s somebody; Something ----s Something
[Eventuality] **shames** [Human] | [Institution]

b. {alter:2} 'become different in some particular way, without permanently losing one's or its former characteristics or essence'
Something ----s; Somebody ----s
[Anything] **alters**

These cases need manual validation to select the correct set of sentence frames and semantic types to be assigned to the verb synset.

4. Towards Conceptual Frames[38]

4.1. Introduction to conceptual frames

A fundamental theory describing the semantic relations between lexical units is **Frame semantics** (Fillmore 1977, 1982, 1985). Frame semantics explains meaning of lexical units through their relation to a semantic frame, a kind of conceptual prerequisite for understanding the meaning of a word (Fillmore 1985: 224), and motivates the semantic frames through the relations of words used in a certain context. In widely accepted terminology it is said that *a lexical unit evokes a semantic frame relating to the specific concept to which it refers*. One of the examples which is often used to illustrate the Frame semantics, is the verb *sell,* a member of the semantic frame **Commercial sell**: the verb *sell* is related with the situation of **Commercial sell** which also involves a seller, a buyer and goods and the specific semantic relations between the seller, the buyer, and the goods.

The semantic resource FrameNet (Fillmore, Baker 2000, 2010) defines semantic frames and organises them in a network. Each semantic frame consists of frame elements: *the various participants, props, and other conceptual roles involved in the semantic representation of a situation* (Fillmore et al. 2003: 359). More than 1,224 semantic frames (1,087 lexical frames) are described so far incorporating 13,640 lexical units, 10,542 frame elements, 10,725 relations of frame elements, and 1,876 relations between semantic frames, i.e., more general frames to more specific ones[39]. For example, the semantic frame **Stimulate emotion** links the core frame elements **Experiencer** (the one who reacts emotionally or psychologically to the Stimulus) and **Stimulus** (the event or entity which brings about the emotional or psychological state of the Experiencer). The semantic frame **Stimulate emotion** organises also the following non-core Frame elements: **Circumstances** (the circumstances under which the Stimulus brings about the Experiencer); **Degree** (the degree to which the Stimulus brings about an emotion in the Experiencer); **Depictive** (a phrase which describes a state of the Experiencer); **Explanation** (the reason why the Stimulus causes the emotion in the Experiencer); **Manner** (the manner in which the Stimulus affects the

---

[38] The chapter was first published as: Koeva, S., T. Dimitrova, V. Stefanova, D. Hristov. Towards Conceptual Frames. – In: *Чуждоезиково обучение*, 46, 6, 2019, pp. 551 – 564. [*Chuzhdoezikovo obuchenie*, 46, 6, 2019, pp. 551 – 564.]

[39] https://framenet.icsi.berkeley.edu/fndrupal/current_status [22 May 2020]

Experiencer); **Means** (the means by which the Stimulus affects the Experiencer); **Result** (the result of the Stimulus affecting the Experiencer); **Time** (the time when the Experiencer has an emotion as caused by the Stimulus)[40].

The information in the FrameNet is derived by annotating sentences in a corpus. The annotation shows the syntactic realisation of frame elements: phrase types and grammatical functions (Ruppenhofer et al. 2016: 7–8). The following annotated sentences are related with the semantic frame **Stimulate emotion** and the verb *irritate*:

> *At theological college, near Oxford,* [*NP* ***the docility of most of the wives of other students=Stimulus***] *IRRITATED* [*N* ***Anna=Experiencer***].
>
> *Nan showed no pleasure at the compliment – in fact Emily seemed to think* [***pro it=Stimulus***] *IRRITATED* [*pro her**=Experiencer***][41].

The presented research aims to specify information about the compatibility of lexical units. For this purpose we need to determine the sets of lexical units with which a Frame element in a particular Semantic frame can be expressed in a sentence. The investigation is limited to the verbal lexical units and to the core frame elements with the type Entity. We call conceptual frame a semantic frame which core frame elements (from the type Entity) are specified for a set of admissible lexical units. Thus, we can reformulate the aims of our research as follows: to enrich FrameNet semantic frames to conceptual frames. In practice this means to describe that the verb *irritate* is combined with an Experiencer which could be **Human** and a Stimulus which could be **Anything**. For example, the Experiencer of the conceptual frame **Stimulate emotion** can be realised among others with the lexical units expressing subordinate concepts of the concept {doctor; doc; physician; MD; Dr.; medico} 'a licensed medical practitioner', such as: *abortionist, allergist, angiologist, gastroenterologist, general practitioner, GP, house physician, resident, resident physician, intern, houseman, medical intern, specialist, medical specialist, surgeon, operating*

---

[40] https://framenet.icsi.berkeley.edu/fndrupal/luIndex [22 May 2020]
Among the lexical units that activate this semantic frame are the verbs: abash, aggravate, aggrieve, agonise, alarm, amaze, anger, annoy, antagonise, appeal, arouse, astonish, astound, baffle, beguile, bewilder, bewitch, boggle, bore, calm, captivate, charm, cheer, comfort, conciliate, confuse, console, crush, dazzle, delight, demolish, depress, destroy, devastate, disappoint, discomfit, disconcert, discourage, dishearten, displease, distress, disturb, embarrass, embitter, enchant, encourage, engage, enrage, entertain, enthrall, exasperate, excite, exhilarate, fascinate, faze, flabbergast, floor, flummox, fluster, frighten, frustrate, fulfill, gall, gladden, grate, gratify, harass, hearten, humiliate, impress, incense, infuriate, interest, intimidate, intrigue, irk, irritate, kill, let down, madden, mollify, mortify, mystify, nettle, nonplus, offend, outrage, pacify, perplex, perturb, petrify, placate, please, puzzle, rankle, rattle, reassure, repel, revolt, rile, sadden, satisfy, scare, shake, shame, shock, shocker.n, sicken, sober, solace, soothe, spook, stagger, startle, stimulate, sting, stir, stun, stupefy, surprise.n, surprise, terrify, thrill, tickle, torment, traumatise, trouble.n, trouble, unnerve, unsettle, upset, vex, vexation.n, worry, wound, wow.

[41] https://framenet.icsi.berkeley.edu/fndrupal/luIndex [The annotated examples are 20 in September 2019.]

surgeon, sawbones, etc. On the other hand, there is no limitation for the lexical units with which the Frame element Stimulus of the Conceptual frame **Stimulate emotion** can be expressed.

The chapter is organised as follows: first we present the motivation for the research and the methods applied combining automatic mapping and manual validation and annotation; later we briefly refer to our previous research and results obtained; the core of the chapter describes the manual validation of automatic mapping and analyses the results.

4.2. Motivation for the research

We exploit two widely known semantic resources: WordNet and PDEV (Pattern Dictionary of English Verbs) patterns with the CPA (Corpus Pattern Analysis) semantic types.

WordNet (Miller 1995; Fellbaum 1998) groups English nouns, verbs, adjectives and adverbs into sets of synonyms (synsets), each expressing a particular concept and thus interchangeable in many contexts. The current number of defined synsets is about 117,000[42]. Synsets are linked by means of conceptual relations. For example, verb synsets are arranged into hierarchies (trees); verbs at the roots of the trees express more abstract concepts while verbs top-down towards the leaves of the trees (called troponyms) express more specific concepts that denotes the manner of doing something (Fellbaum 1990; 2002). The noun synsets are organised into hierarchies (trees) with the super-subordinate relation (hypernymy and hyponymy relations) which links more general concepts to specific ones (with the most abstract concepts being at the root of the tree(s) and most specific concepts at the leaves of the tree(s)) (Miller 1990). Verbs and nouns are grouped into more specific semantic classes (Miller 1990; Fellbaum 1990) describing their general meaning: noun.person, noun.animal, noun.cognition; verb.cognition, verb.change, etc. Each verb synset contains a list of sentence frames illustrating the types of simple sentences in which the verbs in the synset can be used. The WordNet sentence frames represent information for the number of frame elements and brief semantic and syntactic description. For example, the verb *corrupt*, part of the synset {corrupt, pervert, subvert, demoralise, debauch, debase, profane, vitiate, deprave, misdirect} 'corrupt morally or by intemperance or sensuality' is described with the sentence frames: *Somebody ----s somebody* and *Something ----s somebody*.

The main advantages of WordNet for semantic analysis are: a) a large amount of concepts organised in a semantic net; b) grouping of concepts in semantic classes according to their general meaning. A brief comparison between the fundamental approaches underlying the FrameNet and WordNet shows that FrameNet describes the semantic knowledge about the type of event, relation, or entity and the participants (frame elements) involved in it, while WordNet is focused on the description of the semantic relations between concepts.

[42] https://wordnet.princeton.edu [22 May 2020]

The WordNet noun semantic classes can be further specified in order to describe more precisely the groups of words satisfying the verbs' selectional requirements. A very appropriate means for this specification is offered by the PDEV framework (Hanks 2008) within which an ontology of semantic types is developed, where each semantic type refers to properties shared by a set of nouns (Cinkova, Hanks 2010). For example, the PDEV CPA ontology contains semantic types such as [Movie] and [Movie_Part], relevant for the semantic description of verbs like {film, shoot, take} 'make a film or photograph of something'. The PDEV pattern [Human] ----s [Movie] | [Movie_Part][43] represents the CPA semantic types of participants: [Human], {Movie] and [Movie_Part], that are relevant for distinguishing between different senses of these verbs. For a comparison in WordNet, nouns like {movie, film, picture, moving picture, moving-picture show, motion picture, motion-picture show, picture show, pic, flick} 'a form of entertainment that enacts a story by sound and a sequence of images giving the illusion of continuous movement' are classified within the more general semantic class: noun.communication which is also specified for nouns like {dissemination, airing, public exposure, spreading} 'the opening of a subject to widespread discussion and debate'. In some cases the PDEV patterns specify the concrete lexical units if the CPA semantic type consists of one member only, i.e., [Human] ----s cigarette | cigar | pipe.

It is claimed that each PDEV pattern can, in principle, be plugged into a FrameNet semantic frame[44] and some PDEV patterns already contain links to the FrameNet semantic frames. To summarise, PDEV is close to FrameNet in its representation of the semantic knowledge: both semantic resources are based on real corpus examples and encode the core frame elements. Further, the PDEV patterns decode particular (fine-grained) semantic types of frame participants. However, the number of PDEV patterns is relatively small and it could be expected that the CPA ontology could be further enlarged with new members.

4.3. Mapping WordNet with CPA and PDEV patterns

In our previous work, we have presented an effort to enrich WordNet information through the assignment of the CPA ontology to WordNet noun synsets (Koeva et al. 2018 – see chapter 1) and the assignment of PDEV verb patterns to WordNet verb synsets (Koeva et al. 2019 – see chapter 2). This work was performed using a combination of manual setup and automatic assignment based on the setup.

Continuing the effort of expanding WordNet on verb compositionality, PDEV patterns were assigned to WordNet verb synsets. The automatic assignment process used a form of translation, where PDEV patterns were translated to WordNet sentence frames using a set of manually curated translation rules. As PDEV patterns often include alternative and optional participants, some patterns were translated to several sentence frames.

---

[43] https://old.datahub.io/dataset/260d2e3e-1ff3-4c5e-bd11-9238ead750d5 [22 May 2020]
[44] https://nlp.fi.muni.cz/projects/cpa/ [22 May 2020]

4.4. Validation of results of automatic mapping

Our hypothesis is based on the assumption that the literals (as members of a synset) can be interchangeable, thus a pattern which is applicable to one of the literals can be applicable to all other literals. We have manually checked the PDEV patterns assigned, taking into account the following factors: the sense of the verb (synset); the semantic class of the verb synset; the WordNet sentence frames assigned to each synset; the usage examples. We assume that:

(1) The automatically assigned PDEV patterns have to be true with respect to the sense of the verb synset, e.g., the automatically assigned pattern [Human] *hails taxi* | *cab* is not valid for {acclaim; hail; herald} 'praise vociferously'.

(2) The PDEV patterns participants should be defined with the appropriate types in the CPA hierarchy ([Animal], [Dog], [Bird], [Human] are within the [Animate] hierarchy).

(3) The PDEV patterns assigned should be applicable to all the verbs (literals) in a synset – this is hampered by the incomplete data, as not all literals have been assigned patterns from PDEV.

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, {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]. There is variation with respect to prepositions introducing a participant (this usually matches the PP participant in WordNet frames) but for our purpose we assume that patterns with different prepositions are in fact different. In addition, the patterns do not seemingly hold for all the literals, e.g., of all three literals in {accept; consent; go for} 'give an affirmative reply to; respond favorably to', *Somebody ----s something*, there are patterns assigned only to {accept}, while the participant of {consent} is introduced with 'to' (possible frames would be *Somebody ----s PP, Somebody ----s to INFINITIVE*). Thus, in the future new patterns can be added or an existing pattern can be changed (for example, the assigned PDEV pattern [Human 1] *adopts* [Human 2] to the synset {adopt; take in} 'take into one's family' can be modified to cover animate entities: [Human] *adopts* [Animate]). We also observe whether the patterns are valid with respect to the usage examples.

Below, we will discuss a couple of different scenarios.

4.4.1. Invalid PDEV pattern assignments

The issue with overgeneration of patterns is a result of the translation rules where *Someone* was automatically mapped to [Human] and *Something* – to Non-human. Therefore, the verb {blow} 'play or sound a wind instrument' with a WordNet sentence frame *Somebody ----s something* was assigned 9 PDEV patterns ([Human] | [Device] *blows* [Vapour] | [Stuff] | [Physical_Object] *Direction*; [Human] *blows*

([Musical_Instrument]); [Human] | [Road_Vehicle] *blows horn*; [Human] *blows whistle*; [Human] *blows bubble*; [Human] | [Device] *blows hole* *Location*), from which we have validated only two: [Human] *blows* ([Musical_Instrument]) and [Human] | [Road_Vehicle] *blows horn*.

Some of the assigned PDEV patterns have to be deleted as they do not correspond to the sense expressed by the respective synset. To {explode} 'destroy by exploding', two patterns were assigned: [Human] | [Institution] *explodes* [Bomb], and [Human] | [Concept 1] *explodes* [Concept 2], but only the first one is validated.

In other cases, the validation is based also on the usage examples, e.g., with {edit; cut; edit out} 'cut and assemble the components of', two usage examples – "edit film" and "cut recording tape" – is found which correspond to [Human] *edits* [Movie] | [Recording] | [Broadcast], but not to [Human] *edits* [Document].

Another scenario allows for the adaptation of patterns by adding participants if the synset sense allows it. For example, in some cases the sense can be applicable to both [Human] and [Animal] participants, but PDEV patterns with only [Human] were assigned, as in {disembowel; eviscerate; draw} 'remove the entrails of' with the pattern [Human 1] *disembowels* [Human 2] but a usage example of "draw a chicken", which points to further inclusion of [Animal] participant. Another such example is with the five PDEV patterns assigned to {absorb; suck; imbibe; soak up; sop up; suck up; draw; take in; take up} 'take in, also metaphorically', in correspondence with the sentence frame *Something ----s something*, but a usage example (*She drew strength from the minister's words*) points to an additional [Human] participant.

In many cases, wrong PDEV patterns are assigned to the syntactic frame, since the WordNet sentence frame does not specify the type (semantics) of possible participants. Sometimes, this is related to the metaphorical or idiomatic meaning of the verb. The sense of {repair; resort} 'move, travel, or proceed toward some place' was assigned the PDEV patterns: [Human] *repairs* [Artifact]; [Human] | [Institution] *repairs damage to* [Artifact]; [Human] | [Institution] | [Activity] *repairs damage to* [Abstract = Relationship] | [Abstract_Entity]; [Human] *repairs* [Body_Part]. Elsewhere, PDEV patterns different to the assigned ones were applicable to a verb as the metaphoric sense imposes other restrictions. The verb {fire} 'drive out or away by or as if by fire' with usage examples "The soldiers were fired", "Surrender fires the cold skepticism" can only have one PDEV pattern [Human 1] | [Institution] *fires* [Human 2] (*from* [Human_Role]) (and not the other assigned ones such as [Inanimate] | [Stuff] *fires* [Machine]; [Anything] *fires* !Human's! [Psych]; [Anything] *fires* [Human], and [Event] | [Human 1] *fires* [Human 2] | [Emotion] (up)).

The PDEV patterns can also be enriched when the encoding of the participants as Somebody, Something, PP, etc. is not exhaustive, and transitiveness is not systematically followed. For example, the verb {crop; browse; graze; range; pasture} 'feed as in a meadow or pasture' was automatically assigned the pattern [Animal] *browses* (*Location*), but we have added two more patterns: [Animal] *browses on* | *upon* [Stuff], and [Animal] *browses* [Stuff].

4.4.2. Changes to patterns

In 930 synset instances, we have added the respective PDEV patterns as a result of the following assumptions:

1. The WordNet sentence frame participant *Somebody* was automatically matched to the PDEV participant [Human] but it can also be matched to [Human_Group], [Body_Part], [Self]. In some cases, *Somebody* can be matched to [Institution], [Business_Enterprise]. For example, the verb synset {brand; trademark; brandmark} 'mark with a brand or trademark' with frame *Somebody ----s something* was assigned only the PDEV pattern [Human] *brands* [Animal] (*with* [Visible_Feature]), while we added the pattern [Business_Enterprise] *brands* [Artifact] ([Name]). Here, the frame [Human 1] | [Institution] *brands* [Human 2] (*with* [Artifact]) if one assumes that in certain contexts [Human] can be also branded. In some contexts, an entity – Action, Activity, Eventuality – can be associated with a Human, therefore the participants can be expanded, e.g., {afflict} 'cause great unhappiness for; distress', with the frame *Something ----s somebody*, can be expanded with the pattern [Human 1] *afflicts* [Human 2] additionally to [Eventuality] *afflicts* [Human] | [Institution] | [Location].

2. We also assume that prepositional participants – PP in WordNet sentence frames and participants introduced by prepositions such as for, with, on, unto, etc. – and participants expressed via *Direction*, *Location*, etc. may match (mind that *Direction*, *Location*, etc. can be expressed by a PP or an adverbial). For example, when no match was found for the verb {stroll; saunter} 'walk leisurely and with no apparent aim' with a frame *Somebody ----s PP*, we have manually assigned the PDEV pattern [Human] *saunters* *Direction*.

3. A participant cannot be used with some verbs, and they can be removed from the pattern, as with {abort} 'terminate a pregnancy by undergoing an abortion', with the frame *Somebody ----s*, and assigned pattern [Human] | [Animal] | [Fetus] *aborts*. We have assigned an additional pattern: [Human 1] | [Animal 1] *aborts* [Fetus] | [State_of_Affairs], where [State_of_Affairs] = Pregnancy, but [Fetus] is to be removed.

4. There are instances of divergence between the WordNet verb sense and WordNet sentence frames and usage examples which have helped us with the validation. For example, the verb synset {affect; impress; move; strike} 'have an emotional or cognitive impact upon' has only one frame *Something ----s somebody*, but there is *Somebody* in the usage examples "This child impressed me as unusually mature", "This behavior struck me as odd". Therefore, we assigned the patterns: [Eventuality 1] | [Entity 1] *affects* [Human] | [Animal] | [Eventuality 2] | [Entity 2]; [Human] *affects* [Attitude] | [Emotion]; [Anything] *impresses* [Human] | [Institution] (*with* [REFLDET][44] Property]); [Human] *impresses*; and [Artifact] *impresses*.

5. Different literals in a synset may be matched to different PDEV patterns. The synset {tittup; swagger; ruffle; prance; strut; sashay; cock} 'to walk with a lofty proud gait, often in an attempt to impress others' is matched to two patterns with two literals: [Human] *prances* (*Direction*), [Human] *sashays* *Direction*, which differ in respect to the optionality of the *Direction* participant.

6. Subsumption in the CPA hierarchy: The CPA semantic type of a participant may be subsumed under another semantic type, as in the case of the verb synset {brandish; flourish; wave} 'move or swing back and forth' which has the frame *Somebody ----s something* and usage example "She waved her gun" with automatically assigned patterns: [Human] *brandishes* [Weapon]; [Human] *brandishes* [Physical_Object]. In the CPA hierarchy, [Weapon] is a subtype of [Physical_Object], [Inanimate]. Moreover, there is another verb {brandish} 'exhibit aggressively', with the same frame and automatically assigned patterns, and a usage example "brandish a sword", would be more suitable for the [Weapon] participant.

7. PDEV patterns matching with some verbs were not identified at all, as with the verb {crown} 'put an enamel cover on' where the appropriate pattern was neither assigned, nor matched and another pattern was assigned: [Human] *crowns* [Body_Part].

4.4.3 Results in numbers

The manual validation and correction had the following effect:

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

The manually validated PDEV patterns were added to the XML version of the Princeton WordNet verb synsets used for this study and is publicly available under the CC by licence: http://dcl.bas.bg/PWN_PDEV/.

4.5. Analysis of results

The analysis of results show that in most cases the number of participants in the WordNet sentence frames coincide with the number of participants in the PDEV patterns. Consider, for example, the synset {breathe, take a breath, respire, suspire} 'draw air into, and expel out of, the lungs', the root of a WordNet tree with 21 verb synsets. All synsets within the tree are assigned WordNet sentence frames with one or two members: *Somebody ----s*; *Somebody ----s something*, or the combination of both. 15 out of the 23 synsets were assigned PDEV patterns and all except one of them have one or two members. The main difference between encoding semantic and syntactic information in PDEV and in WordNet concerns the granularity of the semantic description with the more granular semantic types in PDEV. The differences can be compared in the following example (synsets and the definitions of their senses are given first (the hierarchy of the hypernyms and hyponyms are shown), the WordNet sentence frames are presented on the left below them, and the PDEV patterns – on the right:

{breathe, take a breath, respire, suspire} ‘draw air into, and expel out of, the lungs’	
Somebody ----s	[Human] | [Animal] ----s
Somebody ----s something	[Human] | [Animal] ----s in [Human] | [Animal] ----s air | dust | gas | [Vapour] (in)
{breathe}	
WordNet synset does not exist	[Fish] breathes (through gills)
{breathe}	
WordNet synset does not exist	[Cetacean] ----s
(respire} ‘breathe easily again, as after exertion or anxiety’	
Somebody ----s	PDEV pattern is not assigned
(choke} ‘breathe with great difficulty, as when experiencing a strong emotion’	
Somebody ----s	[Human] | [Animal] ----s (to death) (on [Physical_Object]) | (on fumes)
{hyperventilate} ‘breathe excessively hard and fast’	
Somebody ----s	PDEV pattern is not assigned
{hiccup, hiccough} ‘breathe spasmodically, and make a sound’	
Somebody ----s	PDEV pattern is not assigned
(sigh, suspire} ‘heave or utter a sigh; breathe deeply and heavily’	
Somebody ----s	PDEV pattern is not assigned
{wheeze} ‘breathe with difficulty’	
Somebody ----s	[Human] ----s
{yawn} ‘utter a yawn, as from lack of oxygen or when one is tired’	
Somebody ----s	[Human] | [Animal] ----s
{snore, saw wood, saw logs} ‘breathe noisily during one’s sleep’	
Somebody ----s	[Human] | [Dog] ----s,
Somebody ----s something	
{exhale, expire, breathe out} ‘expel air’	
Somebody ----s	[Human] | [Animal] ----s
Somebody ----s something	

{snort} ‘make a snorting sound by exhaling hard’	
Somebody ----s	**[Human] | [Animal] ----s**
{blow} ‘exhale hard’	
Somebody ----s	**[Human] ----s (*Direction*)**
Somebody ----s something	
{pant, puff, gasp, heave} ‘breathe noisily, as when one is exhausted’	
Somebody ----s	**PDEV pattern is not assigned**
{puff, huff, chuff} ‘blow hard and loudly’	
Somebody ----s	**[Human] | [Animal] ----s**
[Human] ----s,	
{insufflate} ‘blow or breathe hard on or into’	
Somebody ----s something	**PDEV pattern is not assigned**
{inhale, inspire, breathe in} ‘draw in (air)’	
Somebody ----s	**PDEV pattern is not assigned**
Somebody ----s something	
{aspirate} ‘inhale (air, water, etc.}’	
Somebody ----s something	**[Human] ----s cavity | tube**
	[Human] ----s [Liquid]
{sniff, sniffle} ‘inhale audibly through the nose’	
Somebody ----s	**[Human] | [Animal] ----s**
{snuffle, snivel} ‘snuff up mucus through the nose’	
Somebody ----s	**[Human] ----s**
{snuff} ‘inhale (something) through the nose’	
Somebody ----s something	**[Human] ----s [Artifact]**
{puff, drag, draw} ‘suck in or take (air)’	
Somebody ----s	**[Human] ----s cigarette | cigar | pipe**
Somebody ----s PP	
{huff, snort} ‘inhale recreational drugs’	
Somebody ----s something	**[Human] ----s [Drug]**

The PDEV patterns mediate the mapping of the semantic information in FrameNet and WordNet. The FrameNet information for the core frame elements of the semantic frame **Breathing** corresponds with the semantic information presented at the PDEV patterns and the WordNet sentence frames: **Agent** (the Agent’s breathing causes the motion of the Air) and **Air** (Air is the substance that the Agent causes to move). The lexical units that evoke the semantic frame **Breathing** are: *blow, breathe, exhale, expire, gasp, huff, inhale, inspire, insufflate, pant, puff, respire, sigh, suspire*.

We assume that if a given FrameNet semantic frame is linked through the mediation of the PDEV pattern to a WordNet synset, the corresponding conceptual frame can be enriched with the information for the semantic types of the participants.

Agent = [Human] | [Animal] ----s (Air = air | dust | gas | [Vapour] (in))
Agent = [Fish] ----s (through its gills)

The granular semantic types assigned to WordNet noun synsets define the sets of WordNet literals appropriate to express the core frame elements. For example, the semantic type [Fish], mapped to the synset {fish} 'any of various mostly cold-blooded aquatic vertebrates usually having scales and breathing through gills', is inherited by hyponyms such as: *groundfish, bottom fish, barracouta, snoek, shad, herring, sardine, salmon, trout, whitefish, sea bass, snapper*, etc. Each of these lexical units can express the core frame element Agent in the conceptual frame: **Agent = [Fish] ----s (through its gills),** i.e. ***a salmon | a barracouta | a herring breathes through its gills.***

5. Conclusion

This study describes the process of mapping the information from the Pattern Dictionary of English Verbs (PDEV) to the Princeton WordNet. First, we manually mapped the 253 CPA semantic types to the WordNet concepts, where the hyponyms of a synset to which a CPA semantic type is mapped, inherited not only the respective WordNet semantic primitive but also the CPA semantic type. We presumed this would allow for the establishment of semantic relations between the verb synsets and the noun synsets to which the respective semantic type was mapped. Next, we automatically assigned the PDEV patterns to the WordNet verb synsets to compare PDEV patterns and WordNet sentence frames and manually cross-validated the PDEV patterns and the WordNet hypernyms' sentence frames. Our effort was aimed at making explicit the information for the semantic compatibility of verbs with sets of entities belonging to particular semantic types. We manually validated the annotation of 4,084 patterns, with further assigning 1,568 new patterns and removing 2,815 inappropriate patterns.

Thus, our effort enriched the semantic information in WordNet with more granular semantic classes for the noun synsets and more precise sentence frames for the verb synsets. Some WordNet verb synsets that are not enriched with the PDEV patterns can be automatically assigned such after further analysis of the dependencies of the inheritance of the PDEV patterns in the hypernymy – hyponymy WordNet paths. For others, manual annotation of sentence frames with appropriate semantic types will provide additional information for the compatibility of lexical units. Further, some of the semantic information in FrameNet and WordNet can be combined via the semi-automatic construction of conceptual frames.

References

Bel et al. 2007: Bel, N., S. Espeja, M. Marimon. Automatic acquisition of grammatical types for nouns. – In: *Human Language Technologies 2007: Proceedings of the North American Chapter of the ACL Companion Volume*, pp. 5 – 8.

Bel et al. 2012: Bel, N., L. Romeo, M. Padr. Automatic lexical semantic classification of nouns.– In: *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pp. 1448 – 1455.

Bybee 2010: Bybee, J. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Cinkova, Hanks 2010: Cinkova, S., P. Hanks. Validation of Corpus Pattern Analysis – Assigning Pattern Numbers to Random Verb Samples. <https://nlp.fi.muni.cz/projects/cpa/CPA_valiman.pdf> [01.06.2020]

Crespo, Buitelaar 2008: Crespo, M., P. Buitelaar. Domain-specific English-to-Spanish Translation of FrameNet. – In: *Proceedings of the 6th Language Resources and Evaluation Conference (LREC-08)*, Marrakech, Morocco, pp. 1470 – 1473.

De Cao et al. 2008: De Cao, D., D. Croce, M. Pennacchiotti, R. Basili. Combining Word Sense and Usage for Modeling Frame Semantics. – In: *Proceedings of the Symposium on Semantics in Systems for Text Processing (STEP08)*, Venice, Italy, pp. 85 – 101.

Dorr 1997: Dorr, B.J. Large-scale Dictionary Construction for Foreign Language Tutoring and Interlingual Machine Translation. – In: *Machine Translation*, 12(4), pp. 271 – 322.

Dorr, Olsen 1997: Bonnie J. D., M. B. Olsen. Deriving Verbal and Compositional Lexical Aspect for NLP Applications. – In: *Proceedings of the Eighth Conference of European Chapter of the Association for Computational Linguistics*, pp. 151 – 158.

El Maarouf et al. 2014: El Maarouf, I., J. Bradbury, P. Hanks. PDEV-lemon: A Linked Data Implementation of the Pattern Dictionary of English Verbs based on the Lemon Model. – In: *3rd Workshop on Linked Data in Linguistics: Multilingual Knowledge Resources and Natural Language Processing*, pp. 88 – 93.

Fellbaum 1990: Fellbaum, C. English Verbs as a Semantic Net. – In: *International Journal of Lexicography*, 3 (4), Winter 1990, pp. 278 – 301.

Fellbaum 1998: Fellbaum, C. (Ed.) *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

Fellbaum et al. 2009: Fellbaum, C., A. Osehrson, P. E. Clark. Putting Semantics into WordNets "Morphosemantic" Links. – In: *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, pp. 350 – 358.

Fillmore 1977: Fillmore, C. J. Scenes-and-Frames Semantics. – In: *Linguistic Structures Processing*. Amsterdam, North-Holland, pp. 55 – 81.

Fillmore 1982: Fillmore, C. J. Frame Semantics. – In: The Linguistic Society of Korea (eds.) *Linguistics in the Morning Calm*. Seoul: Hanshin, pp. 111 – 137.

Fillmore 1985: Fillmore, C. J. Frames and the Semantics of Understanding. – In: *Quaderni di Semantica*, vol. 6, pp. 222 – 254.

Fillmore, Baker 2000: Fillmore, C. J., C. F. Baker. FrameNet: Frame Semantics Meets the Corpus. – In: *Poster session at the Annual Meeting of the Linguistic Society of America*.

Fillmore, Baker 2010: Fillmore, C. J., C. F. Baker. A Frames Approach to Semantic Analysis. – In: *Oxford Handbook of Linguistic Analysis.* Oxford: Oxford University Press, pp. 313 – 341.

Fillmore et al. 2003: Fillmore, C. J., C. R. Johnson, M. R. Petruck. Background to FrameNet. – In: *International Journal of Lexicography*, 16 (3), pp. 235 – 250.

Green et al. 2001: Green, R., L. Pearl, B. J. Dorr, P. Resnik. *Lexical Resource Integration across the Syntax-Semantics Interface (No. LAMP-TR-069)*. Maryland University College Park Institute for Advanced Computer Studies.

Hanks 2004: Hanks, P. Corpus pattern analysis. – In: *Proceedings of the 11th EURALEX International Congress*, (EURALEX 2004), Lorient, France, 6–10 July 2004, pp. 87 – 97.

Hanks 2008: Hanks, P. Mapping Meaning onto Use: A Pattern Dictionary of English Verbs. – In: Proceedings of the AACL 2008, Utah. https://slideplayer.com/slide/4126510/> [01.06.2020]

Hanks 2013: Hanks, P. *Lexical Analysis.* Cambridge (MA): MIT Press.

Hanks, Pustejovsky 2005: Hanks, P., J. Pustejovsky. A Pattern Dictionary for Natural Language Processing. – In: *Revue française de linguistique appliquée*, vol. 10 (2), pp. 63 – 82.

Harris 1954: Harris, Z. Distributional structure. – In: *Word*, vol. 10 (23), pp. 146 – 162.

Hindle 1990: Hindle. D. Noun Classification from Predicate-Argument Structures. – In: *Proceedings of the 28th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics*, pp. 268 – 275.

Koeva et al. 2018: Koeva, S., T. Dimitrova, V. Stefanova, D. Hristov. Mapping WordNet Concepts with CPA Ontology. – In: *Proceedings of the 9th Global WordNet Conference (GWC'2018).* Global WordNet Association, Singapore, pp. 70 – 77.

Koeva et al. 2019: Koeva, S., D. Hristov, T. Dimitrova V. Stefanova. Enriching WordNet with Frame Semantics. – In: *Доклади от Международната годишна конференция на Института за български език „Проф. Любомир Андрейчин" (София, 14 – 15 май 2019 година)*. София: Издателство на БАН „Проф. Марин Дринов", 2019, с. 300 – 308. [*Dokladi ot Mezhdunarodnata godishna konferentsiya na Instituta za balgarski ezik „Prof. Lyubomir Andreychin" (Sofiya, 14 – 15 may 2019 godina)*. Sofiya: Izdatelstvo na BAN „Prof. Marin Drinov", 2019, pp. 300 – 308.]

Korhonen 2002: Korhonen, A. Assigning Verbs to Semantic Classes via WordNet. – In: *Proceedings of the 2002 Workshop on Building and Using Semantic Networks*, vol. 11, pp. 1 –7.

Levin 1993: Levin, B. *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press.

McCrae et al. 2012: McCrae, J.P., G.A. de Cea, P. Buitelaar, P. Cimiano, T. Declerck, A. Gómez-Pérez, J. Gracia, L. Hollink, E. Montiel-Ponsoda, D. Spohr, T. Wunner. Interchanging Lexical Resources on the Semantic Web. – In: *Language Resources and Evaluation*, vol. 6 (6), pp. 701 – 709.

Miller 1990: Miller, G. A. Nouns in WordNet: A Lexical Inheritance System. – In: *International Journal of Lexicography*, 3 (4), Winter 1990, pp. 245 – 264.

Miller 1995: Miller, G. A. WordNet: A Lexical Database for English. – In: *Communications of the ACM*, 38 (11), pp. 39 – 41.

Miller et al. 1990: Miller, G. A., R. Beckwith, C. Fellbaum, D. Gross, K.J. Miller. Introduction to WordNet: An Online Lexical Database. – In: *International Journal of Lexicography*, 3 (4), pp. 235 – 244.

Pazienza et al. 2001: Pazienza, M.T., M. Pennacchiotti, F.M. Zanzotto. Mixing WordNet, VerbNet and PropBank for Studying Verb Relations. – In: *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC-2006)*, pp. 1372 – 1377.

Pustejovsky 1995: Pustejovsky, J. *The Generative Lexicon*. Cambridge (MA): MIT Press.

Ruppenhofer et al. 2016: Ruppenhofer, J., M. Ellsworth, M. R. L. Petruck, C. R. Johnson, C. F. Baker, J. Scheffczyk, J. *FrameNet II: Extended Theory and Practice* (Revised November 1, 2016.) <https://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf> [01.06.2020]

Shi, Mihalcea 2005: Shi, L., R. Mihalcea. Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robust Semantic Parsing. – In: Gelbukh, A. (ed.). *Computational Linguistics and Intelligent Text Processing (CICLing 2005). Lecture Notes in Computer Science*, Berlin, Heidelberg: Springer, vol. 3406, pp. 100 – 111.

Tonelli, Pighin 2009: Tonelli, S., D. Pighin. New Features for FrameNet – WordNet Mapping. – In: *Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CONLL-09)*, Boulder, CO, USA, pp. 219 – 227.

Vossen 2002: Vossen, P. *EuroWordNet General Document*. <https://research.vu.nl/ws/portalfiles/portal/77020259/EWNGeneral> [01.06.2020]

Vossen et al. 1998: Vossen, P., G. Hirst, H. Rodriguez, S. Climent, N. Calzolari, A. Roventini, F. Bertagna, A. Alonge, W. Peters. *EuroWordNet: A Multilingual Database with Lexical Semantic*. Kluwer Academic Publisher.

Wierzbicka 1984: Wierzbicka, A. Cups and Mugs: Lexicography and Conceptual Analysis. – In: *Australian Journal of Linguistics*, 4.2 (1984), pp. 205 – 255.

Wierzbicka 1985: Wierzbicka, A. *Lexicography and Conceptual Analysis*. Karoma, Ann Arbor.

Wierzbicka 1986: Wierzbicka, A. What's in a Noun? (or: How do Nouns Differ in Meaning from Adjectives?). – In: *Studies in Language, International Journal sponsored by the Foundation Foundations of Language*, 10.2 (1986), pp. 353 – 389.

Wierzbicka 1988: Wierzbicka, A. *The Semantics of Grammar.* Vol. 18. John Benjamins Publishing.

Svetla Koeva
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria
svetla@dcl.bas.bg

Tsvetana Dimitrova
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria
cvetana@dcl.bas.bg

Valentina Stefanova
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria
valentina@dcl.bas.bg

Dimitar Hristov
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria
dimitar@dcl.bas.bg

This volume features three studies presenting the work on the project *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 Programme (Grant Agreement 10/3 of 14.12.2016).

The authors are researchers at the Department of Computational Linguistics at the Institute for Bulgarian Language, Bulgarian Academy of Sciences. The Department of Computational Linguistics carries out theoretical and applied research in the field of Natural Language Processing. The Department has compiled a large volume of language resources – dictionaries, corpora, as well as a variety of computer applications for language data processing. Among the long-term projects of the Department is the development and enrichment of BulNet – a lexical-semantic network linked to Princeton WordNet and wordnets in many other languages.