Towards Revised System of Verb Wordnet Relations for Polish

Agnieszka Dziob, Maciej Piasecki, Marek Maziarz, Justyna Wieczorek, and Marta Dobrowolska-Pigoń

G4.19 Research Group, Wrocław University of Science and Technology, Wrocław, Poland
{agnieszka.dziob,maciej.piasecki}@pwr.edu.pl

Abstract. A revised model for the description of verbs in plWordNet - a large wordnet of Polish is presented in the paper. The model has originated from the previous one (the versions 2.0 and 3.0) as a combination of simplifications and expansions. The simplifications were motivated by the collected experience of the previous model implementation for almost 18 thousand Polish verbs. The proposed expansions are meant to improve the description of verbs and support provided by plWordNet for various applications.

1 Underestimated Wordnet Part

Verbs are mostly described by fewer relations in wordnets and the verb subnetwork expresses lower density providing less information. In this paper we claim that:

- a richer structure of verb relations can be introduced into a wordnet,
- rich verb relations in a wordnet can supplement valency-based model in a way which is attractive for many applications.

Our goal is to outline a model for the description of verbs in plWordNet which consists of: features (e.g. verb class, aspect) and lexico-semantic relations and is strongly based on the plWordNet general model. We are going to confine the discussion only to the Polish language and plWordNet\(^1\) (plWN - a large wordnet for Polish [30]), but in a perspective enabling more general conclusions. The presented work builds on the previous model proposed for plWN 2.0 [28] and used also in the version 3.0 [30], as well as the experience collected from its implementation for 17,391 verb lemmas. During this work it became clear that the model must be refined in order to improve its applicability in daily wordnet editing practice and consistency of the wordnet editor decisions.

Six verb relations were proposed for Princeton WordNet (PWN) [34]: synonymy, antonymy, inclusive entailment (proper inclusion), troponymy (coextensiveness), cause and presupposition. Later, since PWN 1.5, inclusion and presupposition have been combined into entailment.

GermaNet identified troponymy with hyponymy [22]. PWN’s inclusive entailment was renamed subevent in EWN [1], and ‘entailment’ is only a label for presupposition (backward along the timeline) [15]. Causality is cross-categorial [23]. EWN’s set is similar to GermaNet’s [48]. Hyponymy is PWN’s troponymy. Cause includes PWN’s

\(^1\) http://plwordnet.pwr.edu.pl
presupposition [47]. EWN also introduced near-synonymy (for semantically close co-hyponyms), near-antonymy, as well as cross-categorial synonymy, antonymy and hypernymy.

There are 19 verb relations in plWN 3.0, 6 of them have derivational origin. Table 1 contrasts our selected verb relations with those of PWN, EWN and GermaNet. A list of plWN 4.0 verb relations is presented in Table 2 and discussed in Sec. 3-4.

2 Constitutive Relations and Features

2.1 Synset Definition

The plWordNet model, cf [29; 30] is based on lexical units (LUs, i.e. triples: lemma, part of speech, sense identifier) as basic building blocks. LUs are grouped into synsets on the basis of sharing constitutive relations and features, i.e. two LUs belong to one synset if and only if they share targets of the constitutive relations and values of the constitutive features. Constitutive relations are selected lexico-semantic relations that are enough frequent, express relatively high sharing factor and originate from the tradition of lexicography and wordnets and are relatively well understood, e.g. hyper/hyponymy, holo/meronymy, iterativity, distributivity. Constitutive features constrain the shape of the system of lexico-semantic relations (e.g. aspect or adjective and verb classes) or express some general pragmatic conditions on the use of lexical meanings (e.g. stylistic register). Constitutive features are often referred to in the definitions of the lexico-semantic relations, e.g. hyper/hyponymy is defined only for adjectives of the same class or inter-register synonymy links LUs of non-compatible lexical registers, cf [30].

Summing up this very brief plWN model description, see also [29], lexico-semantic relations are the main means of description. They all can be traced back to the language data. Synset relations are a kind of abbreviations representing sets of constitutive relations held between members of two synsets. The synonymy relation is to some extent derived from the constitutive relations and features that define synsets.

2.2 Aspect

Aspect is one of the fundamental grammatical verb categories and has gained a lot of attention in the literature, cf. e.g. [4; 8; 47]. Comrie describes it as “the internal temporal constituency of the situation” [4]. In Slavic languages it is marked morphologically by affixes and - in some cases - by ablaut, cf. e.g. [3; 4; 19; 24; 35; 49]. According to the most widespread opinion among slavists, grammatical aspect is a way in which a speaker envisages the situation he is speaking about. Thus the imperfective aspect classifies the situation represented by a verb as ongoing, while the perfective aspect informs that the represented situation is delimited in time and described as complete in relation to its elements [24]. Every Polish verb may be either perfective, or imperfective, if only we ignore ca 150 bi-aspectual verbs [31].

Concerning lexical aspect (Aktionsarten) there are many verb pairs with opposition

---

2 The bi-aspectual verbs are ambiguous with respect to the morphological form, and the actual aspect is determined by the context of use [41].
extended beyond imperfective/perfective by additional features of the represented
situation, e.g., the impf. verb *robić* (‘to do, to be doing some activity’) gives birth
not only to its perf. counterpart *zrobić*, but also to many other derivatives formed with
non-empty prefixes, including the delimitative derivative *po-robić* ‘to have been
doing some activity for some time, and to have it finished’ or accumulative derivative
*na-robić się* ‘to work until being exhausted’. plWN covers the grammatical and
lexical aspects jointly in the aspectuality relation. *Pure aspectual pairs* differs only in
the grammatical aspect, while *secondary aspectual pairs* embody also the lexical
aspect [28].
In plWN 3.0 we adopted the theory of grammatical aspect that binds it with the
concepts of telicity/atelicity in such a way that only telic verbs were allowed to form
pure aspectuality pairs. Since telicity depends heavily on syntactic context [46], we
took into account and compared only primary uses of verbs [29], cf [24,35]. However,
the identification of the verb pairs was very laborious, because it required corpus-
based analysis of many examples. Moreover, it was difficult to maintain the
consistency of editors’ decisions. For instance, the pair *gubić* / *zgubić* ‘to lose
something’ looks like an ordinary pure aspectual pair, but the close inspection of a
corpus reveals that *gubić* has the restriction that it rarely occurs with direct objects
expressed with countable NPs in singular (in cursive meaning), which is a clear signal
of atelicity [46; 8]. For many verbs such semantic-syntactic patterns are not so clear
with either telic or atelic contexts being dominant, e.g., the primarily atelic verb
*jeść* ‘to eat’ (being a couple of *zjeść* ‘to eat the apple’). Thus, in the model proposed here, we have resigned from
the telicity condition and based the category on the analysis of definitions and
paraphrases. Thus secondary aspectuality was restricted mainly to those verb pairs
that are overtly linked derivationally by a non-empty affix, mostly a prefix.
Aspect is an intriguing problem for Slavic wordnet designers [13; 20; 21; 36; 37].
Some put pure aspectual pairs into one synset, e.g., BulNet [20] or CzechWordNet
[37], following the model of PWN (in English the similar distinction of progr./perf. is
inflectionally marked, and not derivational in nature). We treat the aspect in an
opposite manner as an important constitutive feature that shapes our verb relations
[29].
The rationale for this decision is following: if we took a closer look into Polish
dictionaries, we could notice that verb definitions (paraphrases) almost always include
superordinate LU (*genus proximum*) in the same aspect as the verb being defined
(*definiendum*). For example, the impf. *robić* ‘to do’ gets the impf. *wykonywać* ‘to execute’ as a *genus*, while the perf. *zrobić* ‘to do’ - the perf. *wykonać* ‘to accomplish’. This is, of course, not a coincidence: you cannot freely substitute in a
corpus a verb of a given aspect with a superordinate verb in an opposite aspect. This
is due to the fact, that although a perf. verb implies its impf. counterpart, the reverse
implication is rarely true (what is in progress not necessarily would be completed
[42]).3 This suggests natural restriction of the hyper-/hyponymy relation only to LUs
of the same aspect. Consequently, if two LUs differ in their constitutive relations they
cannot belong to the same synset, see Sec. 2.1. Bi-aspectual verbs have mostly two
hyponyms: imperfective and perfective and the verb hyponymy tree forks, e.g.

---

3 The exception are parametric pairs, e.g., *rosnąć* impf. - *urosnąć* perf. ‘to grow up’.
mianować ‘to nominate’ can mean both nadać [nadać] ‘to grant [a title]’ but also nadawać [nadawać] ‘to be granting [a title]’.

Pure aspectual pairs that we disunite into separate synsets are interlinked solely by pure aspectuality relation. The meaning differences in the case of secondary aspectuality are even more significant and results also in additional lexico-semantic relations between the pair elements.

Aspect is also correlated with other verb relations, e.g. multiplicativity:iterativity is observed only for imperfective verbs and tests for meronymy are constrained by aspect, see Sec. 3.

2.3 Verb Classes

In plWN 3.0 verbs were organised into a hierarchical system of classes derived from the proposal of Vendler [45] and Laskowski [24] for Polish. Classes and subclasses were represented by the top level verb hypernym synsets. However, the classification was fine-grained, multi-level and appeared to be difficult in consistent use. Moreover, the representation of classes as top synsets made shaping the structure a challenging task, e.g. all co-hyponyms had to belong to one single fine-grained class. For plWN 4.0 we decided to significantly simplify the classification and its representation. While still being close to Vendler [45], we use only two top classes: state and dynamic verbs.

State verbs are imperfective atelic and durative verbs that represent situations which are stable and do not include any change during their time span. State verbs can represent [24]: 1) localisation or position in space, (e.g. leżeć ‘to lie’,stać ‘to stand’, znajdować się ‘to be in some place’, otaczać ‘to surround’), 2) possession of permanent material features like mass, temperature, state of matter (e.g. ważyć ‘to weigh’, mierzyć ‘to be of particular size’, świecić ‘to produce light’), 3) relationships between entities, both material and non-material, like to be a part of something, to belong to someone/something, to be in a space (e.g. ograniczać ‘to delimit’, równać się ‘to equal to’, składać się ‘to consist of’), 4) mental states, emotional, sense experience (e.g. odczuwać, czuć both meaning ‘to feel’, wiedzieć ‘to know’, doświadczać ‘to experience’, kochać ‘to love’, widzieć ‘to know’), 5) other static situations (e.g. żyć ‘to live’, spać ‘to sleep’, spodziewać się dziecka ‘to be pregnant, lit. to expect a child’).

State verbs are all imperfect and are linked with their perfective counterparts - dynamic verbs - by the secondary aspectuality relation.

Dynamic verbs include all former auxiliary classes of plWN 3.0 [28]: 1) distributive verbs (to do something by many agents or in relation to many objects, e.g. powykańczać się ‘to get exhausted, finished, plural sub., about many things’, powylamywać ‘to be breaking off many things/parts’), 2) accumulative (to do something to such an extent that it is enough, e.g. najeść się ‘to be eating until an agent is full’, spracować się ‘to be working so long that it is enough’), 3) perdurative (to be doing something during particular or limited time, e.g. przepracować ‘to be working during some time’, przewiegetować ‘to be vegetating during a particular time’), 4) delimitative (to be doing or to happening for some time or to some extent, e.g. pobiegać ‘to be running a little’, popadać ‘to be raining a little’).
All other dynamic verbs are grouped in 5) *action* verbs: a) all perfective and bi-aspectual, b) imperfective derivatives of accumulative, delimitative, perdurative, and distributive verbs (representing changing situations), c) imperfective derivatives of semelfactive verbs (i.e. representing punctual or instantaneous events), e.g. *mrugać*<sub>impf</sub> (from *mrugnąć*<sub>perf</sub>) 'to flicker', representing multiple changes, d) imperfective causative verbs (expressing occurrence of a new situation), e) processive (expressing gradual transition to a new state), f) inchoative (representing introductory phase of a new situation) and g) limitative (representing ending of some situation), and also g) all other imperfective verbs that represent situation changing due to actions of entities involved or with respect to them, e.g. *iść* 'to walk<sub>imp</sub>', *biegać* 'to run<sub>imp</sub>', *plakać* 'to cry<sub>imp</sub>.'

The proposed system is much less fine-grained and more shallow - two main classes and 5 subtypes - than the one of plWN 3.0 (9 main classes, 4 auxiliary classes). In plWN 3.0 only some of these verb classes had direct influence on the system of verb relations. Thus the former auxiliary classes became subclasses of the dynamic verbs. Their direct influence on the relation structure was limited only to distributive verbs related to the distributivity relation, see Sec. 4.3.

The main reasons for distinguishing state verbs were the cross-categorial synset relation called *state* [28], see Sec. 3.4, imperfective aspect of state verbs and also the secondary aspectuality which links state verbs with their dynamic counterparts.

### 3 Synset Relations

#### 3.1 Hypernymy Hierarchical Structures

In PWN *troponymy* relation was used instead hyponymy for verbs [9], and described as “to V1 is to V2 in some particular manner”. It is called also “a manner relation” and links temporally co-extensive verbs. However, EWN has opted for verb hypernymy which can be adapted to the manner relation [48]. GermaNet identified troponymy with hyponymy [22]. Maziarz et al. [29] compared the ways of defining nouns and verbs in dictionaries and argued for the use of hyper/hyponymy relations for verbs in plWN 3.0. It links verbs of the same semantic class, aspect and compatible register.

Derwojedowa et al. [5] argued that there is place for both relations: troponymy and hyponymy in plWN due to the specific morphology of Polish verbs. Many verbs are derived by prefixes from their derivational bases that seem to be their hypernyms in the same time. Troponymy could be a relation signalling a derivational association between a more specific and more general verb. However, starting from plWN 2.0 such association started to be described by secondary aspectuality relation.

Derwojedowa et al. [5] and Piasecki et al. [38] noticed the necessity of having two verb relations expressing specification of the meaning as a result of the rich semantics of verb prefixes in Polish. In GermaNet these two tendencies are combined [15; 16] by putting on the upper levels of the verb hierarchy synsets that exemplify the meaning of verb prefixes. In Polish, the semantics of the verb prefixes is more

---

4 Semelfactive verbs as perfective verbs are included into actions, see the point a).
complex and its more in depth description is planned for further research. pWN 3.0 verb hierarchy resembles the one from PWN 3.1 where top levels are populated by artificial synsets (non-lexicalised)\(^5\) that are similar to 40 “top-level synsets” in PWN [11]. We plan to rebuild it in such a way that the upper levels of the hierarchy will be also synsets grouping hyponymy branches, but constructed in a bottom-up direction. The reconstruction work will be started from the most specific verbs in the bottom levels and proceed to upper levels. The top synsets will result from the language data and potential needs for more comprehensible organisation of the tree in the case of artificial synsets. As in the case of nouns, we will not combine all verb synsets into one single-rooted tree. Semantic class and aspect are kept as constitutive features that determine the hypernyomic trees for two reasons. First, the simplified set of classes is easier in use for wordnet editors and next the class assignment facilitates identification of hypernyms. Second, state verbs are all imperfective, while their perfective derivates are in fact secondary aspectuality pairs and express additional semantic difference. Division into classes supports consistent organisation of the verb synsets.

### 3.2 Associations between Situations

Besides hyper/hyponymy, the proposed verb constitutive relations (i.e. synset relations) can be divided into 3 groups. The first group includes lexico-semantic relations expressing associations between situations that are related to the timeline to some extent. It includes: presupposition and preceding (backward relations), meronymy (co-occurrence of two situations), inchoativity (beginning of a situation), causality (resulting in a situation) and processuality (a state resulting). They are only slightly amended since pWN 3.0. If we observe a situation represented by a verb from the perspective of a point \(t_0\) in the present time, then causality, processuality and inchoativity refers to a point \(t_1\) in the future resulting somehow from the verb situation, presupposition and preceding refer to a point \(t_{-1}\) in the past, as the verb situation is influenced by what had happened around \(t_{-1}\), while meronymy refers to the present time period and a co-occurring situation, cf [28]. In pWN 3.0 model all these relations had many subtypes related to properties of LUs linked (e.g. aspect). According to our experience, definitions and substitution tests have been simplified by parameterizing them by the feature values. WordnetLoom [39] application has been extended in such a way that on the basis of the aspect attribute of a LU, an appropriate version of a test is generated and presented to the editor.

<table>
<thead>
<tr>
<th>PWN</th>
<th>GermaNet</th>
<th>EWN</th>
<th>pWN</th>
</tr>
</thead>
<tbody>
<tr>
<td>synonymy (V-V)</td>
<td>synonymy (V-V)</td>
<td>synonymy (V-V)</td>
<td>synonymy (V-V)</td>
</tr>
<tr>
<td>antonymy (V-V)</td>
<td>antonymy (V-V)</td>
<td>antonymy (V-V, N, Adj, Adv)</td>
<td>antonymy (V-V)</td>
</tr>
<tr>
<td>troponomy (V-V)</td>
<td>hyponymy (V-V)</td>
<td>hyponymy (V-V)</td>
<td>hyponymy (V-V)</td>
</tr>
</tbody>
</table>

\(^5\) Representing very fine-grained division into semantic classes.
Except meronymy (V-V) and inchoativity (V-V,N) all other relations can link verbs to any other part of speech. The system of 6 situation association relations can seem to be complex. In PWN, most of them are covered by entailment with cause as the only exception [9]. A more fine grained description was proposed in EWN [48], see Table 1. However, some relations are signalled derivationally, more fine-grained types helps to write more precise definitions and substitution tests and all of them are relatively frequent, see statistics from plWN 3.0 in Tab. 2. Due to the above mentioned problems with the model of plWN 3.0 we can expect that the presented numbers are not complete.

**Inchoativity** [28] links: V-V,N, see Tab. 2, where the first verb represents an initial phase of a situation represented by the second element [24]. Inchoativity is signalled derivationally, i.e. the second verb LU is a derivational basis for the first, e.g. rozpłakać się ~to start crying' -inch.→ plakać `to be crying', but due to its semantic regularity was expanded to LU pairs not linked derivationally, e.g. {aktywować się 1 'to activate itself'} -inch.→ {funkcjonować 1 'to function', działać 3 'to work', chodzić 4 ~to work, pracować 3 'to work', ...}, in all these cases aktywować się 1 means 'to start X-ing'. In pairs with nouns, the noun names a situation whose beginning is represented by the verb, e.g. rozpłakać się 1 means `~to start crying' -inch.→ płacz 'cry'.

**Processuality** [28] can be described as ‘to become or to be becoming’. It is often signalled by a derivational association, but it was also expanded to a synset relation, e.g. czerwienić 1= `to be becoming’ czerwony 1 `red’ or zmieniać się 1 `to be changing itself/yourself’= to be becoming ‘inny 1 ‘different’.

**Causality** is used in PWN [33] to link only verbs, in EWN is cross-categorial. pWN causality is defined for dynamic verbs (as in EWN), and the effect, intentional or unintentional, can be represented by a dynamic or static verb, as well as any other PoS [28]. For instance ogrzać 1 `to warm up’ means: 1) `to cause smth’ zagrząć się 1 `to have become warm by itself’, 2) `to cause heat [cieplo 1]’, 3) `to cause that smth. is warm [ciepły 1]’, 4) `to cause that it is warm’ [spowodować, że jest ciepło 1]’. EWN cause relation encompases also cause non-factial relations. In pWN this type of relationship is expressed by preceding relation and the stricter cases of necessity by presupposition [28]. Presupposition is close to logical presupposition and informs about the necessity of earlier occurrence of some situation, e.g. żyć 1 `to live'
Umrzeć 1 'to die', because if someone has died, he had to live earlier, the same if he has not died; thus umrzeć 1 presupposes żyć 1.

**Preceding** relation represents a possibility that one situation happens before the other one, e.g. siedzieć 'to sit', stać 'to stand' ← prec. położyć się 'to have laid down', if someone has laid down, he could earlier sit or stand, but both precede the former.

Because during construction of plWN 3.0 we could observe problems with consistent selection between causality and presupposition/preceding, we propose here to introduce two subtypes of both: identical (e.g. martwy 'dead' ← pres. zmartwychwstać 'to resurrect'; żona 'a wife', męż 'a husband' ← prec. rozwiesić się 'to divorce') and non-identical subjects (e.g. zabić 'to kill' ← pres. żyć 'to live'; minimum 'a minimum', maksimum 'a maximum' ← prec. przekroczyć 'to overcome'). The latter subtype of presupposition and preceding resembles causality, but in opposite direction, and we assume that both can be use in parallel but in opposite directions.

### Table 2. Verb relations in plWordNet 4.0.

<table>
<thead>
<tr>
<th>Relation</th>
<th>POSs</th>
<th>Example</th>
<th>No(3.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>inter-register</td>
<td>V-V</td>
<td>pieprzyć [vulgar] 'to speak nonsense' → mówić 'to speak'</td>
<td>2016</td>
</tr>
<tr>
<td>synonymy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aspectuality</td>
<td>V-V</td>
<td>napisać 'to write' - pisać 'to write'</td>
<td>26558</td>
</tr>
<tr>
<td>- pure</td>
<td></td>
<td>popisać 'to write a little' - pisać 'to write'</td>
<td></td>
</tr>
<tr>
<td>- secondary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hyponymy</td>
<td>V-V</td>
<td>pływać 'to swim' → przemieszczać się 'to move'</td>
<td>22680</td>
</tr>
<tr>
<td>inchoativity</td>
<td>V-V, N</td>
<td>rozplakać się 'to start crying' → płakać 'to cry'</td>
<td>403</td>
</tr>
<tr>
<td>processuality</td>
<td>V-N, Adj, Adv</td>
<td>owdzić 'to become a widow or widower' → wdowa 'a widow', wdowiec 'a widower'</td>
<td>729</td>
</tr>
<tr>
<td>causality</td>
<td>V-N, Adj, Adv</td>
<td>suszyć 'to cause that smth. is drying' → schnić 'to dry'</td>
<td>1773</td>
</tr>
<tr>
<td>presupposition</td>
<td>V-V, N, Adj, Adv</td>
<td>umrzeć 'to die' presupposes żyć 'to live'</td>
<td>167</td>
</tr>
<tr>
<td>preceding</td>
<td>V-V, N, Adj, Adv</td>
<td>siedzieć 'to sit' precedes wstać 'to stand up' (but wstać does not presuppose siedzieć)</td>
<td>169</td>
</tr>
<tr>
<td>meronymy</td>
<td>V-V</td>
<td>przędykać 'to swallow' is an integral part of situation jeść 'to eat'</td>
<td>2829</td>
</tr>
<tr>
<td>holonymy</td>
<td>V-V</td>
<td>jeść 'to eat' is a typical situation including przędykać 'to swallow'</td>
<td>2888</td>
</tr>
<tr>
<td>multiplicativity</td>
<td>V-V</td>
<td>grywać 'to play a little from time to time' → grać 'to play'</td>
<td>542</td>
</tr>
<tr>
<td>- iterativity</td>
<td></td>
<td>pokraść 'to steal much' → ukraść 'to steal'</td>
<td></td>
</tr>
<tr>
<td>- distributivity</td>
<td></td>
<td>czekać 'to wait' → gotowy 'ready'</td>
<td>93</td>
</tr>
<tr>
<td>state antonymy</td>
<td>V-V, N, Adj, Adv</td>
<td>oddać 'to give back' - zatrzymać 'to keep'</td>
<td>2348</td>
</tr>
<tr>
<td>- complementary</td>
<td></td>
<td>zwiększać 'to increase' - zmniejszać 'to reduce'</td>
<td></td>
</tr>
<tr>
<td>- proper</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>converseness</td>
<td>V-V</td>
<td>sprzedać 'to sell' - kupić 'to buy'</td>
<td>112</td>
</tr>
<tr>
<td>role inclusion</td>
<td>V-N</td>
<td>oceniać 'to shade' ← cień 'a shade'</td>
<td>1357</td>
</tr>
<tr>
<td>- subject</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Meronymy (holonymy) was defined in PWN and EWN only for nouns. Because plWN verb hyponymy does not cover cases in which a situation is an element of a larger, more general situation, verb meronymy was introduced to express such associations [28] with two subtypes of sub-situation and accompanying situation. As this distinction was not very clear in many cases, for plWN 4.0, we have removed both subtypes and put emphasis on the necessary simultaneous co-occurrence of two situations. A verb meronym represents a more narrow situation, e.g. chrapać 2 'to snore' -mero→ spać 1 'to sleep'. Holonomy is a reverse relation, but is not automatically defined.

3.3 Multiplicativity

Multiplicativity relation describes repetition of some state or activity, and performing an activity by many subjects or on many objects. The first subtype is called iterativity, and the second distributivity, cf [28; 49]. Iterativity links imperfective verbs, while distributivity perfective. Both are expanded also to verbs pairs that come from the same root, but which are not direct derivates. Multiplicativity relation is more specific than hyponymy and excludes it. Multiplicative verbs are linked with their derivational bases of different aspects by the lexical relation (non-synset) of secondary aspectuality (e.g. nakupować 1 'to buy perf many things' -sec. asp.- kupować 1 'to buy imp').

3.4 Cross-categorial relations

Constitutive relations are the primary means of defining synsets and we are still looking for possibilities to increase the number of their instances and to increase the network’s density. Moreover, cross-categorial relations are valuable for many wordnet applications, including WSD. State relation, introduced in plWN 2.0, in plWN 4.0 is restricted to the state verbs. It represents being in some state, e.g. ograniczać 3 'to delimitimp' means ‘to be a border [granica 1]’, czerwienić się 1 ‘to be red’ -state→ czerwony 1 ‘red’, czuwać ‘to stay awake’-state→ świadomie 'exalted'
We propose four new cross-categorial relations in plWN 4.0. **Manner** relation describes a verb by a link to an adverb describing a manner in which the activity is performed:

Jeżeli ktoś lub coś X-wał(o), to znaczy, że Z-wał(o) Y-owo. (gdzie X jest hyponimem Z).

‘If smn/smth has X, it means that he/it has Z Y[ADV]’, where X is a hyponym of Z’ e.g. podgotować ‘to cook perf a little’ → trochę [Adv] ‘little [Adv]’;

**Circumstance** describes a verb by referring to an adverbial realised by a simple prepositional phrase and links the verb to the noun which is the semantic head of the prepositional phrase:

Jeżeli ktoś lub coś X-wał(o), to znaczy, że Z-wał(o) Y[Prep]. (gdzie X jest hip. Z.)

‘If smn/smth has X, it means that has Z Y[Prep]’ e.g., dopływać ‘→ swimię to some point/place’ -circum. → brzeg ‘a bank’, where the phrase do brzegu ‘to the bank’ is an unexpressed element.

Subject and object are conceptually generalised from the morpho-semantic role relation (see Sec. 5.2) It was motivated by functional grammar associating semantics with syntax [18], e.g. subject and object relations in [6]. They characterise semantic roles of the semantic subject and object [43]. Inclusion of predicate-argument roles as definition elements can result in their proliferation and subjectivity in adding them. That is why, we limited the number of instances per verb to 3. However, we aim at adding only very characteristic links, so smaller number, better. Such links must originate from meaning definitions.

Subject: Jeżeli mowa o tym, że ktoś lub coś X-ował(o), to ten ktoś lub to coś pewnie był(o) Y-kiem. ‘If we say that smn/smth X, then this smn/smth is certainly Y’ e.g. mucić ‘→ moo’ → krowa ‘a cow’.

Object: Jeżeli mowa o tym, że X-owo kogoś lub coś, to ten ktoś lub to coś pewnie był(o) Y-kiem. ‘If we say that smn/smth was X, then this smn/smth had to be certainly Y’ e.g. wzuwać ‘→ put shoes on’ but ‘a shoe’.

4 Lexical Relations

Relations that do not fulfill requirements for constitutive relations are described on the level of LUs, not synsets. This type includes first of all semantic oppositions (like in PWN and most other wordnets), but also a large number of derivationally motivated relations (that is common for Slavic wordnets), called also morpho-semantic relations.

4.1 Oppositions

Opposition relations have not changed since plWN 3.0, cf [28; 38]: **pure antonymy**, **complementary antonymy**, and **converseness**. All are restricted to verbs. Converseness, considered in PWN, but not included, cf [9] links verbs representing the same situation, but from different viewpoints and with the reverse positions of arguments in the semantic valency frame, e.g. sprzedać 1 ‘to sell’ - kupić 1 ‘to buy’.
4.2 Morpho-semantic relations

Derivational relations mostly express some lexical meaning change, but derivational morphemes are often very ambiguous, cf [40]. The idea of morpho-semantic relations was proposed, e.g. [12] to overtly describe these semantic links, and used in wordnets, e.g. [48]. Morpho-semantic relations are especially numerous in Slavic wordnets, e.g. [20]. In EWN relations of this type does not need to be strictly derivationally motivated that is similar to our expansion of some morpho-semantic relations, namely: inchoativity, multiplicativity, processuality, and state to synset relations, see Sec. 4. In addition a lexico-semantic relation of role cf. [14] was introduced in plWN 2.0 [28], and has been preserved unchanged. Role is described only if a noun is derived from a verb (role inclusion describes derivation of a verb from a noun), and has 7 subtypes (see Tab. 2). In plWN 4.0 role will be registered only for imperfective verbs, as most perfective verbs are derived from imperfective. Such a solution should help to improve consistency in adding role instances and limits its artificial proliferation.

5 Potential Applications and Conclusions

The modified and extended system of verb relations will lead to a denser network and richer description of verbs. Cross-categorial links are especially important for WSD in relation to all PoSs. WSD methods based on activation spreading across the wordnet network express increasing performance with the increasing number of words from the given text context that are connected in the network.

Relations describing associations between situations in Sec. 4.2 form a kind of a system. On the basis of our former experience, we are convinced that they can be defined with good consistency and they are relatively frequent. Moreover, the current description, see Tab. 2, is definitely not complete. So, such relations should facilitate calculating graph-based text similarity, recognition of textual entailment, text semantic relations or paraphrases, as well as even some forms of reasoning.

The former, plWN 3.0, verb model was complex, and some problems have appeared in its application in wordnet development. Some elements of the model (e.g. classes and relations) were not interpreted in a way consistent enough, many relation subtypes have not been described with good coverage. Our applications of plWN (e.g. WSD, text similarity) revealed a need for further increase of the verb network density, especially in relation to cross-categorial links, e.g. there are significant differences in WSD quality depending on the density of the local wordnet subnetworks around different verbs.

The proposed modifications in the plWN verb model are more an evolution than revolution. In most cases we try to simplify the model, but in a careful way, from the perspective of the wordnet development process. We try to firmly base new proposed relations on the analysis of the language data. References to the intuition of linguists are inevitable, but not the core of the definitions. We try to perceive plWN as a multi-purpose, semi-formal description of the Polish lexical semantics systems, but necessarily built on a large scale and focused on range of applications in NLP.
Acknowledgments

This work was co-financed as a part of the investment in the CLARIN-PL research infrastructure (www.clarin-pl.eu) funded by the Polish Ministry of Science and Higher Education.

References


