TSVETANA DIMITROVA$^A$ & VALENTINA STEFANOVA$^B$

Institute for Bulgarian Language, Bulgarian Academy of Sciences, Sofia, Bulgaria

$^A$cvetana@dcl.bas.bg; $^B$valentina@dcl.bas.bg

THE SEMANTIC CLASSIFICATION OF ADJECTIVES IN THE BULGARIAN WORDNET: TOWARDS A MULTICLASS APPROACH

Abstract

The paper presents an attempt at semantic classification of adjectives in the Bulgarian wordnet. Although designed for the Bulgarian wordnet, the classification can be applied to other wordnets which are developed in parallel to the Princeton WordNet. The classification relies on information that is already available in WordNet from other synsets (noun, verb, and other adjective synsets) that are linked to the adjective synsets via lexical-semantic relations — including their semantic classes, as well as definitions and usage examples. The first stage of the work was already presented at the workshop “Challenges for WordNets” within the conference ‘Language, Data and Knowledge 2017’. The continuation of the effort as described in this article, covers a proposal for introducing additional semantic classes to the adjective synsets (if applicable).

Keywords: WordNet; Bulgarian language; lexical semantics; semantic classification; adjectives

1 Introduction

Each synset in the Princeton WordNet is assigned a semantic class$^1$ (Fellbaum, Osherson, & Clark, 2009; Miller et al., 1993). However, unlike noun and verb synsets which are subjected to elaborate semantic classifications — nouns are organized into 25 semantic classes such as noun.person, noun.animal, noun.plant, noun.process, noun.event, noun.act, noun.state, noun.group, noun.location, noun.time, noun.vehicle, etc., and verbs — into 15 semantic classes — verb.stative,

---

$^1$Terms such as semantic primitive or prime, semantic class, and semantic label have been used, sometimes interchangeably, throughout the literature on WordNet. Despite the intention declaring that each semantic class will be treated as a ‘unique beginner’ of a separate hierarchy (Miller, Beckwith, Fellbaum, Gross, & Miller, 1993, p. 16), further practice went contrary to it. Here, we adhere to the term ‘semantic class’ and, occasionally, ‘(semantic) label’.
verb.communication, verb.change, verb.cognition, verb.contact, verb.perception, verb.cause, etc.,
the classification of adjective is too sparse — only three labels are applied to the adjective synsets — adj.all (mainly) for descriptive adjectives, adj.pert for pertainyms, and adj.ppl for adjectival participles. In consideration of this, we propose a more detailed semantic classification of adjectives to be introduced into the Bulgarian wordnet. Our effort — which is a work-in-progress — draws upon a number of existing classifications proposed in the linguistic literature for Bulgarian as well as the semantic classification of adjectives in German as being applied to the Wordnet for German (GermaNet, see Hamp & Feldweg, 1997).

The approach relies on the relational network of the Princeton WordNet because the Bulgarian Wordnet (Koeva, 2008; Koen, 2014) largely keeps the original structure of the Princeton WordNet (while introducing closed class words such as pronouns, prepositions, conjunctions, particles, and interjections, as outlined in (Koeva, Tinchev, & Mihov, 2004).

The paper is organized as follows. In Section 2, we briefly discuss our starting point (namely, the classifications of adjectives in other wordnets, as well as in the linguistic literature on Bulgarian). Section 3 presents in more detail the implemented classification of adjectives covering: the ways in which we have identified candidate adjective synsets (which, at this point, include only adjectives marked with the generic label adj.all) in the Bulgarian wordnet that will be further classified by taking into account the relevant information from nouns, verbs and other adjectives linked via lexi-semantic relations; an overview of the co-occurrence of adjective and noun classes within the synsets’ trees (within the verification process); and a detailed review of the classification at present — here, we exemplify only adjective synsets labeled with more than one semantic class label following the comments on the first stage of our effort described in Stefanova and Dimitrova (2017). Some directions for future work are given in Section 4.

2 Adjectives in WordNet

The concepts in WordNet (called synonym sets) are nodes linked to each other via lexi-semantic relations (edges) that encode relations between the synsets such as hypernymy/hyponymy, meronymy (membership, partiality, etc.), antonymy, synonymy, similarity, derivation, as well as a number of morpho-semantic relations between nouns and verbs such as agent, undergoer, instrument, event, result, location, material, etc.

In addition, each WordNet synset is classified by a semantic class (Miller et al., 1993; Fellbaum et al., 2009) — noun synsets are organized into 25 semantic classes, while verb synsets are labeled under 15 semantic classes. Adjectives (in the Princeton WordNet) are classified into two larger classes: descriptive and relational adjectives; plus adjectival participles (Fellbaum, Gross, & Miller, 1993). They are accordingly labeled — adj.all, adj.pert and adj.ppl — and are organized into separate non-intersecting structures — each class is linked to other synsets via different sets of relations. Descriptive adjectives (labeled by the semantic class adj.all) are organized into clusters based on similarity of meaning (synonymy) and binary opposition (antonymy). Relational adjectives (labeled adj.pert) are (derivationally) related and subsequently linked to the synset which contains (as a literal) their source noun. Although the Bulgarian wordnet follows the structure of the Princeton WordNet, there are a number of language-specific concepts and features that are to fit in. For example, relational adjectives in Bulgarian often have no morphological equivalent in English (which often uses nouns as modifiers, as in dog food in English — kucheshka hran in Bulgarian) — when (solely) introduced into the Bulgarian wordnet, they are linked to the respective noun via the pertainym relation. Adjectival participles are marked as adj.ppl and are related via participle relation to synsets containing the verbs they are derived from.

Thus, adjectives are organized via a set of relations encoding their properties of attribution, antonymy, similarity, derivation; fuzzynymy and thematic category (which have been added in the EuroWordNet (Vossen, 2002), etc.; some relations are specific for one of the two classes (attribute, similarity, fuzzynymy — for adj.all; pertainym — for adj.pert; participle — for adj.ppl) (Dimitrova
The semantic classification of adjectives in the Bulgarian Wordnet: Towards a multiclass approach

Tsvetana Dimitrova & Valentina Stefanova

Other wordnets such as the Wordnet for German (GermaNet; Hamp & Feldweg, 1997), the Polish WordNet (plWordNet; Maziarz, Szpakowicz, & Piasecki, 2012), and the WordNet for Russian (RussNet; Azarova & Sinopalnikova, 2004) have introduced other classifications of adjectives.

Adjectives in the Polish WordNet (plWordNet) are categorized into four classes based on derivativity: deadjectival (derived from other adjectives), deverbal adjectives, quality adjectives (mostly denominal that are often derived from nouns via the attribute relation when express the values of an attribute), and relational adjectives. In plWordNet 2.0, a set of lexical-semantic relations was introduced to connect adjectives to the other parts of the wordnet as discussed in (Maziarz et al., 2012). The WordNet for Russian (RussNet; Azarova & Sinopalnikova, 2004) sticks to the traditional terms and organizes adjectives into four classes: descriptive adjectives that are subdivided into adjectives expressing size, and qualitative adjectives; relational adjectives; pronominal adjectives (derived from pronominal base); and ordinal adjectives (numerals).

As we have opted for a semantic classification, however, our approach partly channels the one that has been adopted by the WordNet for German (GermaNet). It is based on a classification proposed by Hundsnurscher and Splett (1982) which employs the modification property of the adjective—a (modifying) adjective is (semantically) linked to a certain (modified) noun to form a separate semantic entity. Thus, thirteen semantic fields (Perzeption, Ort, Zeit, Bewegung, Substanz, natPhaenomen, Koerper, Gefuehl, Geist, Verhalten, Gesellschaft — social, Gesellschaft — social quantity, Relation) and three general semantic fields (Allgemein, Pertonym, Privativa) are hierarchically organized and further divided according to sub-features expressed by 70 (sub)classes organized around a specific feature.

Some of these classes were also found in the linguistic literature on Bulgarian where, traditionally, the adjectives are divided into two large classes—qualitative and relational adjectives (roughly, in parallel to adj.all and adj.pert classes in WordNet) — and are analyzed as a class of dependent lexical items whose semantic and syntactic properties are fully realized only in relation to the noun they modify Radeva (В. Радева, 1991). Thus, even when an adjective expresses a property of being related to an object or an event (as is the case with relational adjectives), it expresses a relational property of another object or event that manifests in a certain way, to a certain degree or in relation to a certain internal property of the modified object or event.

A few works on adjectives also claim that there is no clear-cut division between qualitative and relational adjectives (Barbolova, 1997; Граматика на съвременния български книжовен език, 1983; Кушаров, 2007; Б. Радева, 2011; В. Радева, 1991; Спасова, 1959).

In Radeva (В. Радева, 1991), an adjective is claimed to denote a property that is “permanently inherent for the entity and is attributed to it in its entirety. This is also the reason to define the adjective as part-of-speech whose denotative function is realized through its connection to the noun.” According to the author, however, there is no clear distinction between qualitative and relational adjectives. Qualitative adjectives can be derived from nouns and refer to an attribute property of the defined entity (expressed by the noun) through its relationship to another entity. The property qualifies and characterizes the entity expressed by the noun from which they are derived (e.g., pitiful – pity, malicious – malice, etc. — examples are from Radeva; В. Радева, 1991). Hence, the hypothesis is that the adjective expresses one-sided relationship with the entity denoted by the motivating noun; the property is perceived as permanent and inherent (В. Радева, 1991, p. 161).

In another study Radeva (В. Радева, 2011) proposes a combined class of qualitative-relational adjectives due to the lack of clear distinction between the two traditional classes, on the one hand, and, on the other, because many relational adjectives acquire qualitative meaning when combined with certain nouns or even in consideration with the source word. Earlier, Trifonova (Трифонова, 1982, p. 169) claims that qualitative and relational adjectives can be part of a general grammar–semantic type while maintaining their relative autonomy.

2The translation from the works in Bulgarian has been adapted.
Radeva (Б. Радева, 2011) also states that the relation of the adjective to the motivating noun determines the class and the core meaning of the adjective, but an essential part can be attributed to the derivation, as well — specific qualitative sense is realized with relational adjectives when combined with a limited class of nouns (kucheshka hrana ‘dog food’ – kucheshki zhivot ‘doglike life’ – kucheshka viarnost ‘dog faithfulness’ – kucheshki stud ‘doggy cold’ (Б. Радева, 2011, p. 29).

In the next section, we will discuss the semantic classes that we have formulated. The first stage of our effort where a set of adjective synsets (over 2,500) were manually classified, was already described in an earlier paper (Stefanova & Dimitrova, 2017). Here, we give additional details — covering all the semantic classes (with the number of classified adjectives being raised to 3,110) — and present our further effort for introducing additional semantic classes with some of the adjectives.

3 Semantic classification

We have formulated 17 semantic classes of adjectives covering: social and community affiliations (social-related, Gesellschaft in GermaNet (GN)); place or location; local time (time, Zeit in GN); weather (natPhaenomen in GN); physical characteristics (body, Koerper in GN); movement (motion, Bewegung in GN); knowledge (cognition, Geist in GN); attitude (relation, Relation in GN); feeling (Gefuuehl in GN); behavior (Verhalten in GN). Three of the semantic classes in GermaNet are too general and include adjectives that can be considered attributes: General (Allgemein), Pertainyms (Pertonym), Privative (Privativa).

The classification needs to comply with the classifications of nouns and verbs that have been already implemented through introducing semantic classes to noun and verb synsets. For example, we expand the class of physical properties of (adj.body with adjectives associated with physical properties of animals (coating, fur, tail, etc.) and plants (flowering, roots, etc.). The substance class is additionally re-organized to cover ingredients; while material class is intended for man-made objects. Additionally, we have introduced classes for adjectives denoting a state (of a person or an object), a causing phenomenon or trigger of change of state, and adjectives that express quality properties of animate and inanimate objects.

3.1 Attribution and verification

First, we have classified a set of adjectives manually by applying only one semantic class to an adjective synset while taking into account the information that is already available in wordnet such as the semantic class of noun and verb synsets linked via lexico-semantic relations to the adjective to be classified, the semantic class of other adjectives in the synset structure, and definitions and usage examples.

To determine the semantic class of an adjective, we search in the database of the Bulgarian wordnet using regular expressions (the Bulgarian wordnet can be locally accessed through the Hydra system3 — the regular expressions and the modal language are outlined in Rizov, 2014). We have retrieved information through:


(a disease that leads to physical changes in / on the body of a person’, ‘body’, ‘appearance’ – adj.body. For example, the query \[\text{sem\_class('adj.all')&definition('in time')}] \) returns a synset such as \{antecedent:1, anterior:1, subsequent:1, precedent:1, previous:1\})

(ii) Usage examples – they may exemplify reference to persons, objects and / or phenomena. There are adjectives that express characteristics that are only found with persons (humans) but not with objects (such as hair and skin color, feelings, etc.). Although the adjectives of the semantic class adj.behavior and adj.feeling can be modifiers to activities of a person, they are still related to persons and not objects, whereas the semantic class adj.substance expresses properties of an object.

(iii) Structure of the synset — if the semantic class of an adjective cannot be easily determined, the structure of the synset is reviewed. For example, \{izgladnial:1; pregladnial:1; umrial ot glad:1; osvirepial ot glad:1\} / \{famished:1; ravenous:2; sharp-set:1; starved:1; esurient:3\} is linked to \{ostarviavan:e:1; osvirepavan:e:1\} / \{edacity:2; esurience:2; ravenousness:1; voracity:2; voraciousness:3\} which is classified as noun.state, but its hypernym is \{glad:1\}, \{hunger:5; hungriness:3\} — a physiological state manifested by an acute need of food. Therefore, the adjective can be classified as adj.state but the state is related to body — adj.body. A set of semantic classes were automatically introduced through extraction from the database of synsets of derivationally related nouns and verbs (and adjectives, later on). Classified adjectives were checked and edited if the definition and usage examples did not match the selected class because there were discrepancies between the synsets.

(iv) Similarity / synonymy — direct or indirect relations to other adjectives that have been already classified: the query \[\text{lang('bg')&pos('a')&<eng\_derivative>$s]\] returns synsets \($s\) in Bulgarian \(\text{lang('bg')}\) of adjectives \(\text{pos ('a')}\) linked to other synset(s) via derivative relation: \{razhdasht se:1\} / \{nascent:1\} is linked to \{razhdane:3; rozhdenie:1\} / \{birth:2; nativity:1; nascency:1; nascence:1\} of the semantic class noun.event; the query \[\text{lang('bg')&sem\_class('adj.behavior')&<antonym>$s]\] returns synsets in Bulgarian classified as adj.behavior which have an antonym: \{strakhliv:1; plashliv:2; uplashen:2\} / \{cowardly:1; fearful:4\} — antonym \{khrom-bur:1; muzhestven:1; druzhoven:1\} classified as adj.behavior.

### 3.2 Semantic classes: interdependence

The adjective classes comply with the semantic classes of nouns and verbs — either the ones directly linked to the adjective or the ones that are indirectly linked via other adjectives in the synset tree. Adjectives are linked to noun and verb synsets via lexico-semantic relations: \{ has_attribute \}; \{ eng\_derivative \}; \{ category\_domain \}; \{ usage\_domain \} — linking the synsets in a topical class. Table 1 contains an overview of the semantic classes of nouns to which a set of 3,110 adjective synsets are directly linked (and not indirectly, i.e., via other adjective synsets directly linked to noun synsets).\(^4\)

\(^4\)The table covers all semantic classes that have been formulated within the first stage, as a first layer (i.e., additional classes — within the multiclass effort — are not included); only the summed up numbers (given in the second column) are exhaustive but not the others.
<table>
<thead>
<tr>
<th>adj.class</th>
<th>noun.class</th>
<th>derivative, noun.class</th>
<th>category, noun.class</th>
</tr>
</thead>
<tbody>
<tr>
<td>body</td>
<td>all</td>
<td>17: attribute: 16(...)</td>
<td>58: state: 19, body: 13, attribute: 10, animal: 8, person: 2, plant: 2 (...)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>51: cognition: 45(...)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>perception</td>
<td>all</td>
<td>22: attribute: 22</td>
<td>54: attribute: 34, food: 10, cognition: 6(...)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>quantity</td>
<td>all</td>
<td>21: attribute: 20(...)</td>
<td>20: quantity: 6, attribute: 6, relation: 3(...)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>cognition</td>
<td>all</td>
<td>19: attribute: 9, cognition: 8(...)</td>
<td>62: cognition: 30, state: 14, person: 10, communication: 6 (...)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>40: cognition: 35(...)</td>
</tr>
<tr>
<td>time</td>
<td>all</td>
<td>21: attribute: 17(...)</td>
<td>18: time: 8, attribute: 5, state: 2 (...)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9: cognition: 9(...)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>social</td>
<td>all</td>
<td>16: attribute: 6(...)</td>
<td>28: person: 13, attribute: 5, cognition: 3, group: 2, location: 2 (...)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7: cognition: 3, act: 3 (...)</td>
</tr>
<tr>
<td>location</td>
<td>all</td>
<td>14: attribute: 8, location: 6(...)</td>
<td>13: location: 6, object: 3, artifact: 2 (...)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>state</td>
<td>all</td>
<td>17: state: 12, attribute: 5</td>
<td>28: state: 14, time: 4, person: 3, attribute: 2 (...)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9: cognition: 3, act: 3(...)</td>
</tr>
<tr>
<td>relation</td>
<td>all</td>
<td>2</td>
<td>34: relation: 15, person: 5 (...)</td>
</tr>
<tr>
<td>substance</td>
<td>all</td>
<td>2</td>
<td>27: substance: 17, attribute: 4, object: 3 (...)</td>
</tr>
<tr>
<td>motion</td>
<td>all</td>
<td>3: 2: state, 1: attribute</td>
<td>5: attribute: 1, verb: change: 3</td>
</tr>
<tr>
<td>weather</td>
<td>all</td>
<td>—</td>
<td>9: phenomenon: 5, artifact, substance: 3 (...)</td>
</tr>
<tr>
<td>material</td>
<td>all</td>
<td>—</td>
<td>1: artifact</td>
</tr>
</tbody>
</table>

Table 1: Distribution of adjective classes and noun classes.
An adjective synset is linked to other adjective synsets via relations of antonymy (\(<\) antonym \(>)\), and similarity (via relations \(\text{also\_see}\) — to link semantically related synsets, and \(\text{similar\_to}\) — for semantic similarity between focal and satellite synsets with close referential meaning). We assume that semantically related adjectives can be classified into the same semantic class but there are a number of exceptions as exemplified in Table 2 for the combinations in synset trees of \textit{adj.behavior}, \textit{adj.feeling}, \textit{adj.perception} and \textit{adj.cognition}.

<table>
<thead>
<tr>
<th>adj.class</th>
<th>adj.class</th>
<th>similar_to</th>
<th>also_see</th>
</tr>
</thead>
<tbody>
<tr>
<td>adj.cognition</td>
<td>adj.behavior</td>
<td>26</td>
<td>5</td>
</tr>
<tr>
<td>adj.cognition</td>
<td>adj.cognition</td>
<td>101</td>
<td>40</td>
</tr>
<tr>
<td>adj.behavior</td>
<td>adj.feeling</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>adj.behavior</td>
<td>adj.behavior</td>
<td>117</td>
<td>62</td>
</tr>
<tr>
<td>adj.perception</td>
<td>adj.perception</td>
<td>230</td>
<td>29</td>
</tr>
<tr>
<td>adj.feeling</td>
<td>adj.feeling</td>
<td>75</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 2: Combination of adjective classes in adjective synset trees.

<table>
<thead>
<tr>
<th>adj.class: number</th>
<th>adj.class</th>
<th>adj.class</th>
<th>adj.class</th>
<th>adj.class</th>
<th>referent</th>
</tr>
</thead>
<tbody>
<tr>
<td>cognition:159</td>
<td>state:77</td>
<td>behavior:22</td>
<td>feeling:8</td>
<td>relation:9</td>
<td>person:74</td>
</tr>
<tr>
<td>time:156</td>
<td>quantity (duration):46</td>
<td>motion:6</td>
<td>relation:20</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>feeling:94</td>
<td>behavior:21</td>
<td>state:21</td>
<td>cognition:12</td>
<td>relation, quality:1</td>
<td>mostly person</td>
</tr>
</tbody>
</table>

Table 3: Combination of adjective classes in adjective synset trees.

An adjective may express a property of an entity but the definition can be worded to include references to different entities such as human, animal, plant (for \textit{adj.body}); or reference to behavior, feeling, cognition; etc. For example, \{perceptive:1\} ‘having the ability to perceive or understand; keen in discernment’ can be used both for human abilities and for actions attributed to human abilities, and the synset is linked to nouns classified as \textit{noun.attribute}, \textit{noun.cognition}, \textit{noun.feeling}, as well as to \textit{verb.cognition}. The adjective \{consumptive:2\} — which is classified as \textit{adj.state} — with the definition ‘afflicted with or associated with pulmonary tuberculosis’, and examples ‘a consumptive patient’ (a state of a person — and reference to \textit{adj.body}) and ‘a consumptive cough’ (a hint at \textit{adj.cause} — ‘a cough caused by tuberculosis’). There are more than one way to
resolve these issues — we may either split a synset, or classify it under more than one semantic class, and we have opted for the latest, as for now.

Additional semantic classes have been applied manually only to a set of adjectives that were automatically extracted from the database based on whether there is a (direct) relation to a noun synset or an (indirect) relation (no further than one step) via a related adjective in the synset tree (if there is a noun synset linked to an adjective synset linked via a \((\text{similar\_to})\) or \((\text{also\_see})\) relation to the adjective at hand). Some numbers for the adjective synsets that received additional semantic classes, are given in Table 3 (synsets with the most variety are given; more than two semantic classes can be applied). The classes \textit{person, animal, plant, food} refer to the referent of the (modified) noun.

The example below\(^5\) illustrates a complex adjective synset which has been classified according to its definition but whose synset tree involves adjectives for perception, state, cognition, feeling.

Example:

\{bezchuvstven:1; nechuvstvitelen:2\} / \{a: insensitive:3\}
\begin{itemize}
  \item \textit{adj.body, adj.perception}
  \item ‘incapable of physical sensation'
  \item \textit{also see}: \{nesçuznavasht:1; neosçuznavasht:1\} / \{unaware:1; incognizant:1\} \textit{adj.cognition}
  \item \textit{also see}: \{bezçuznatevlen:1; bezchuvstven:2\ v bezçuznavanie:1\} / \{unconscious:2\} \textit{adj.state}
  \item \textit{also see}: \{nechuvstvitelen:1\} / \{insensitive:1\} \textit{adj.body, adj.perception}
  \item \textit{also see}: bg \{bezchuvstven:3; nechuvstvitelen:3\} / \{insensitive:2\} \textit{adj.feeling}
  \item \textit{antonym}: \{chuvstvitelen:9\} / \{sensible:4; sensitive:3\} \textit{adj.perception}
  \item \textit{eng\_derivative}: \{n: bezchuvstvenost:9; nevçuzpriemchivost:1; nechuvstvitelnost:3\} / \{insensibility:2\} \textit{noun.cognition}
  \item \textit{similar\_to}: \{anesteziran:1; obezbolen:1\} / \{anesthetic:1; anaesthetic:1\} \textit{adj.body, adj.perception}
  \item \textit{similar\_to}: \{vtzepenen:1; iztrçupal:1; vkochanen:1; vkochaniasal:1\} / \{asleep:5; benumbed:2; numb:2\} \textit{adj.body, adj.state, adj.perception}
\end{itemize}

We cannot argue for a hierarchy among the semantic classes applied, though, for practical reasons, the ones that have been applied first (and can be checked at: \url{http://dcl.bas.bg/bulnet/}) are given first. For some adjectives for feeling, cognition, behavior (behaviour is modeled by emotive and cognitive patterns; expressed emotions and exercised cognitive abilities are related to perception), the classification proved rather tricky (cf. the synset \{bezchuvstven:1; nechuvstvitelen:2\} / \{a: insensitive:3\} ‘incapable of physical sensation’ and the tree with related synsets which combine synsets of all of these semantic classes). For others — such as those referring to body parts — we have adopted a (practical) approach that the adjective expressing a property of a body or location on a body, (and classified as \textit{adj.body}) will come first because being a body part can be an intrinsic relation expressed by the adjective (this is true also for the hair and eye color in cases where an \textit{adj.perception} can also be applied).

### 3.3 Semantic classes

The attempted semantic classification of the adjectives as applied to the Bulgarian Wordnet combines some of the classes outlined above (mostly adopted from the GermaNet classification) plus some information from noun and verb classes (following the Princeton WordNet classes — the

\(^5\)Examples throughout the paper include: the synset in Bulgarian; the synset in English; the definition in English; the semantic class, plus additional semantic classes, if applicable; other synsets linked via lexico-semantic relation to the adjective synset at hand, with semantic classes given in italic. Beware that only one semantic class of the adjective synset — if given — can be viewed in the database of the Bulgarian wordnet which can be checked at: \url{http://dcl.bas.bg/bulnet/}. There is a parallel view with Bulgarian vs. English, the parallel synsets are marked by a red arrow beside the synset (for description of the viewer and its features, cf. Rizov, Dimitrova, & Barbu Mititelu, 2016). Examples throughout the paper do not give the whole synset tree.
The semantic classification of adjectives in the Bulgarian Wordnet: Towards a multiclass approach

Semantic classes have been previously validated and a number of changes have been introduced into the Bulgarian wordnet (as well as in the local database of the Princeton WordNet) — the effort and results are described in Koeva, Leseva, Stoyanova, Dimitrova, & Todorova, 2016.6 As already stated, the continuation of the work is to introduce additional semantic classes to adjectives if: (i) they can express relations and properties that can be classified into more than one semantic class (e.g., the adjectives expressing relation to time (adj.time) can express a position in time or a duration of an act); (ii) the adjective expresses characteristics that can be attributed to a referent of specific class (described in Stefanova & Dimitrova, 2017).

In the examples below, we include only adjectives with two or more semantic classes (plus additional classes of the referent of the noun the adjective can modify (person, animal, plant) — these are applied to adjective classes with a variety of modified noun’s referents.

(1) Adj.body: these are adjectives that express physical characteristics of humans, animals, and plants, and the nature of the possessor of these characteristics is given as an additional semantic class. An adjective may refer to a property that can be considered under another semantic class but if it can be attributed to body part, it is still classified as adj.body.

(1a)
{liuspestopopashat:1} / {scaly-tailed:1}
adj.body, animal
‘having a scaly tail’
similar_to: {opashat:1} / {caudate:1; caudated:1}
adj.body, animal

(1b)
{zaden:6; gruben:4} / {back:7; hind:1; hinder:1}
adj.body, adj.location, animal, person
‘located at or near the back of an animal’
similar_to: {zaden:1} / {posterior:2} adj.body, adj.location, animal, person

(1c) {tufest:1} / {floccose:1}
adj.body, plant
‘(of plants) having tufts of soft woolly hairs’
similar_to: {kosmat:1; okosmen:1; kosmest:1}, {hairy:2; haired:1; hirsute:1} adj.body, animal, person
category_domain: {rastenie:1} / {plant:1; flora:1; plant life:1} noun.plant

(2) Adj.cognition: adjectives denoting cognitive processes and contents and expressing cognitive abilities of a person or entities resulting from the cognitive activities of a person; in most cases, the class of the referent of the modified noun is not marked (for example, most of the adjectives related to grammatical or other linguistics abilities are classified as adj.cognition only; the class of the referent, however, is included in adjectives specifically referring to the state of the person, as in the example (2a)).

(2a)
{obrazovan:1} / {educated:1}
adj.cognition, adj.state, person
‘possessing an education (especially having more than average knowledge)’
also_see: {tsivilizovan:1} / {civilized:1; civilised:1} adj.social
also_see: {intelektualen:1} / {intellectual:1} adj.cognition

6Changed data is described and can be downloaded at: http://dcl.bas.bg/en/wordnetMSRs/.
(2b) \{obsesiven:1; natrapliv:2\} / \{obsessional:1; obsessive:1\}
adj.cognition, adj.behavior
‘characterized by or constituting an obsession’

eng._derivative: \{kompulsia:1\} / \{compulsion:3; obsession:2\} noun.motive
eng._derivative: \{vmaniachavam se:1; vmaniacha se:1\} / \{obsess:2\} verb.emotive
eng._derivative: \{maniia:5; obsebvane:3\} / \{obsession:1; fixation:3\} noun.cognition
similar_to: \{nevrotichen:2; nerven:3; psikhonevrotichen:1\} / \{neurotic:1; psychoneurotic:1\}
adj.cognition, adj.state, person

(3) Adj.feeling: this class covers adjectives for feelings and emotions of a person or related entities. Here, the additional semantic classes can denote behavior, cognition, etc.

(3a) \{sranezhliv:1; sramliv:1; stestnitelen:1\} / \{shy:2\}
adj.feeling, adj.behavior
‘wary and distrustful; disposed to avoid persons or things’
similar_to: \{predpazliv:4; vnimatelen:5\} / \{wary:1\} adj.behavior

(3b) \{bezchuvsten:11\} / \{unfeeling:1\}
adj.feeling, adj.state
‘devoid of feeling or sensation’
usage: shy of strangers
similar_to: \{bezchuvsten:10; neodusheven:5\} / \{insentient:1; insensate:1\} adj.state

(4) Adj.behavior: these are adjectives that express behavior, behavioral symptoms, etc. that are usually characteristic to person (and more rarely, to animal). Additional semantic classes may include adjectives referring to feeling, cognition, etc.

(4a) \{nereligiozen:1\} / \{irreligious:1\}
adj.behavior, adj.cognition, person
‘hostile or indifferent to religion’
also_see: \{neblagochestiv:1\} / \{impious:1\} adj.behavior
antonym: \{religiozen:2; viarvasht:2\} / \{religious:2\} adj.behavior, adj.cognition, person

(4b) \{mil:3; sırdechen:3\} / \{kind3:1\}
adj.behavior, adj.quality
‘having or showing a tender and considerate and helpful nature; used especially of persons and their behavior’
also_see: \{sgovorchiv:1\} / \{accommodating:1; accommodative:1\} adj.behavior

(5) Adj.perception: adjectives for seeing (color), hearing (voice, sound), and perception (taste, sense, physical sensation, etc.) and, rarely, estimation (liking/disliking, etc.) which can be included as an additional semantic class.

(5a) \{mandarinen:1\} / \{tangerine:1\}
adj.perception, adj.relation
‘of a strong reddish orange color’
similar_to: \{khromatichen:1; otsveten:2\} / \{chromatic:1\} adj.perception

(5b)
\{svetül:1\} / \{fair:3; fairish:1\}
adj.perception, adj.body
‘(used of hair or skin) pale or light-colored’
similar_to: \{rus:1; rusokos:1; svetlokos:1\} / \{blond:1; blonde:1; light-haired:1\} adj.body, adj.perception

(6) Adj.time: adjectives expressing age, historical period, succession in time, longevity, occurrence in a specific time period. Additional semantic classes may express motion adj.motion, duration adj.quantity, etc. In addition, the adjective may express a property related to the age of a person, an animal or a plant.

(6a)
\{bienalen:1; dvugodishen:2\} / \{biennial:2; biyearly:4\}
adj.time
‘occurring every second year’
similar_to: \{periodichen:1; periodicheski:1\} / \{periodic:2; periodical:1\} adj.time

(6b)
\{dvugodishen:1\} / \{biennial:1; two-year:1\}
adj.time, adj.quantity
‘having a life cycle lasting two seasons’
\{antonym\}: \{ednogodishen:1\} / \{annual:1; one-year:1\} adj.time, adj.quantity
\{antonym\}: \{mnogogodishen:3\} / perennial:2 adj.time, adj.quantity
\{category_domain\}: \{botanika:1\} / \{botany:1; phytology:1\} noun.cognition
\{eng_derivative\}: \{dvugodishno rastenie:1\} / \{biennial:3\} noun.plant

(6c)
\{dvugodishen:3\} / \{two-year-old:1\}
adj.time
‘two years of age’
similar_to: \{mlad:1; nezrial:3\} / \{young:5; immature:5\} adj.time, adj.state, person, animal

(7) Adj.location: adjectives expressing spatial properties, placement, succession in space, etc. Examples with additional semantic classes are quite few (some were already classified as semantically referring to another property or entity, and adj.location is the additional class).

(7a)
\{tropicheski:1\} / \{tropical:3\}
adj.location
‘relating to or situated in or characteristic of the tropics (the region on either side of the equator)’
\{eng_derivative\}: \{tropik:1\}, \{Torrid Zone:1; tropical zone:1; tropics:1\} noun.location
similar_to: \{ekvatorialen:1\} / \{equatorial:1\} adj.location

(8) Adj.motion: adjectives related to manners of motion (vehicle, speed, etc.); no additional classes have been applied yet — only the referent of the (modified) noun can be marked.

(8a)
\{baven:3\} / \{'slow:8'\}
adj.motion
‘not moving quickly; taking a comparatively long time’

\textit{similar\_to}: \{bavnodvizhesht\ se:1; bavnopodvizhen:1\} / \{slow-moving:2\} \textit{adj.motion}
\textit{antonym}: \{b˘ urz:2\} / \{fast:5\} \textit{adj.motion}

(9) \textit{Adj.social}: adjectives that express relations resulting from social norms and principles or concern entities or phenomena that are part of the social structure (incl. religion, ideology, marriage, etc.). Additional semantic classes may relate some characteristics of the social norms or phenomena (ideology — cognition; age — \textit{adj.time}).

(9a) \{desen:2\} / \{right:25\}
\textit{adj.social, adj.cognition}
‘of or belonging to the political or intellectual right’
\textit{also\_see}: \{konservativen:1\} / \{conservative:2\} \textit{adj.social, adj.cognition}
\textit{antonym}: \{liav:2; levicharski:2\} / \{left:5\} \textit{adj.social, adj.cognition}

(10) \textit{Adj.substance}: these are adjectives expressing relation to substance which can be related to different entities (plant, animal, etc.)

(10a) \{d˘ urvenist:1\} / \{ligneous:1\}
\textit{adj.substance, plant}
‘consisting of or containing lignin or xylem’
\textit{category\_domain}: \{botanika:1\} / \{botany:1; phytology:1\} \textit{noun.cognition}
\textit{eng\_derivative}: \{lignin:1\} / \{lignin:1\} \textit{noun.substance}
\textit{similar\_to}: \{d˘ urvesen:1; d˘ urvesinen:2\} / \{woody:3\} \textit{adj.substance, adj.material, plant}

(10b) \{mesen:2\} / \{meaty:2\}
\textit{adj.substance, animal, food}
‘like or containing meat’
\textit{antonym}: \{bezmesen:1\} / \{meatless:1\} \textit{adj.substance, animal, food}
\textit{eng\_derivative}: \{meso:2\} / \{meat:2\} \textit{noun.food}

(11) \textit{Adj.material}: adjectives expressing materials used for production of man-made objects.

(11a) \{d˘ urven:1\} / \{wooden:1\}
\textit{adj.material, plant}
‘made or consisting of (entirely or in part) or employing wood’
\textit{similar\_to}: \{d˘ urvesen:1; d˘ urvesinen:2\} / \{woody:3\} \textit{adj.substance, adj.material, plant}

(12) \textit{Adj.weather}: adjectives related to climate conditions (a limited number).

(12a) \{tropichen:1\}, \{tropical:1\}
\textit{adj.weather, adj.location}
‘of weather or climate; hot and humid as in the tropics’
\textit{eng\_derivative}: \{tropic:1\} / \{tropic:3\} \textit{noun.location}
\textit{similar\_to}: \{goresht:3; nagoreshten:1; paresht:1\} / \{hot:9\} \textit{adj.perception}

(13) \textit{Adj.quantity}: adjectives expressing quantity, size, degree, range, etc. Adjectives that denote only quantity as in (13a), rarely have another semantic class but the label can be found as an
additional class (13b) with adjectives referring to the size of a body or a body part, a portion of time, etc.

(13a) {mnogobroen:2}, {numerous:1; legion:1}

adj.quantity

‘amounting to a large indefinite number’

eng_derivative: {mnogobroynost:1; mnogochislenost:1; mnozhestevenost:3}, {numerousness:1; numerosity:1; multiplicity:2} noun.attribute

{has_value}: {mnogo:6} / {many:1} noun.quantity

(13b) {dvugodishen:1} / {biennial:1; two-year:1}

adj.time, adj.quantity

‘having a life cycle lasting two seasons’

category_domain: {botanika:1}, {botany:1; phytology:1} noun.cognition

eng_derivative: {dvugodishno rastenie:1}, {biennial:3} noun.plant

(14) Adj.state: adjectives expressing a state of a person or an entity which is more or less stable for a period of time but can be subjected to change (physical, cognitive, etc.). Additional classes may express the nature of the entity affected.

(14a) {tuberkulozen:2} / {tubercular:1}

adj.state, adj.body

‘constituting or afflicted with or caused by tuberculosis or the tubercle bacillus’

eng_derivative: {tuberkuloza:1} / {tuberculosis:1; TB:3; T.B.:1} noun.state

eng_derivative: {tuberkul:3} / {tubercle:3} noun.state

similar_to: {bolen:3; zle:3} / {ill:8; sick:7} adj.state, adj.body

(15) Adj.cause: adjectives expressing abilities relating to change of state. An additional class can refer to a change in the physical state (as in (15a) or in the psychological state (as in (15b).

(15a) {virulenten:1; bolestotvoren:2} / {virulent:1}

adj.cause, adj.body

‘infectious; having the ability to cause disease’

antonym: {avirulenten:1; nebolestotvoren:1; apatogenen:1} / {avirulent:1} adj.cause, adj.body

eng_derivative: {virulentnost:1} / {virulence:2; virulency:2} noun.attribute

similar_to: {sm˘ urtonosen:3} / {deadly:1} adj.cause, adj.body

(15b) {strashen:2} / {atrocious:1; frightful:1; horrifying:1; horrible:1; ugly:1}

adj.cause, adj.feeling

‘provoking horror’

eng_derivative: {strashnost:1; strakhovitost:3} / {frightfulness:1} noun.attribute

similar_to: {trevozhen:3; obezpokoitelen:1; zastrashitelen:1; bezpokoiasht:2; obezpokoiasht:1} / {alarming:1} adj.cause, adj.feeling

(16) Adj.quality: adjectives expressing a property that is considered more or less characteristic of the entity it characterizes. This is a rather broad class and it needs additional classification depending on the nature of the property it refers to (e.g., the adjective in (16a) refers to both
physical and moral characteristics which can be subjected to a change still — hence, the additional adj.state class.

(16a) 
\{zelen; zlachen:1\} / \{verdant:1\}
 adj.quality, adj.state
 ‘conducive to or characteristic of physical or moral well-being’
similar_to: \{obilen:2; izobilen:2; izobilstvasht:1\} / \{abundant:1\} adj.quantity

(17) Adj.relation: adjectives denoting an explicit relation to an entity such as possession, purpose, function, composition, similarity, etc. Examples for additional semantic classes here involve adjectives which express also social characteristics — related to social structures, the social situation of a person, but not directly including him as an object; cognition — concerning but not directly involving cognitive structures or properties.

(17a) 
\{mužki:1\} / \{male:3\}
 adj.relation, person
 ‘for or pertaining to or composed of men or boys’
eng_derivative: \{maleness:1; masculinity:2\} noun.attribute
eng_derivative: \{male:5; male person:1\} noun.person

(17b) 
\{kucheshki:2\} / \{doglike:1\}
 adj.relation,animal
 ‘resembling a dog; especially in devotion’
similar_to: \{loialen:2; veren:9\} / \{loyal:2\} adj.behavior

The classification outlined above has been applied mostly to check the relevance of the semantic classes. However, there is a significant work to be done — both on classification of more adjectives and on further (sub)dividing of some classes and determining the interdependence between others (e.g., adj.state and adj.quality). There is also a possibility to enrich the present set of semantic classes, e.g., with ones expressing size, degree, age — as subclasses of time, adj.quantity, adj.body, etc.; possession; belonging to certain group — subclasses of adj.social; taste, color, smell, physical sensations — subclasses for adj.perception, etc.).

4 Future work

As a continuation of our project, we plan to use further heuristics to extract and (semi-)automatically label the adjectives incl. comparative ‘bag-of-words’ representation of definitions (and usage examples, etc.); similarity of synset terms, relative tree location.

Another direction of research that will help for verification of the semantic classes involves taking into account possible collocations between adjectives, on the one hand, and nouns, on the other, on the basis of the existing derivational relations between them in consideration with certain semantic classes (e.g., whether noun.person and noun.animal often co-occur with adj.body; noun.person — with adj.feeling; noun.food — with adj.perception; etc.). This is based on the observation that two words can co-occur only if they are semantically compatible, thus collocations can be considered an intrinsic part of a word’s meaning.
The semantic classification of adjectives in the Bulgarian Wordnet: Towards a multiclass approach

References


Радева, В. (2011). Относителни прилагателни имена в съвременния български език (семантика, граматика) (Дисертация). Софийски университет “Св. Климент Охридски”.


Радева, В. (2011). Относителни прилагателни имена в съвременния български език (семантика, граматика) (Дисертация). Софийски университет “Св. Климент Охридски”.

References (transliteration)


Acknowledgment

Part of this work was carried out within the project “Semantic Classification of the Adjectives in the Bulgarian WordNet” of the Institute for Bulgarian Language which is supported under the Program for Career Development of Young Scientists (2016) at the Bulgarian Academy of Sciences.

We would also like to thank the three anonymous reviewers and the participants in the workshop Challenges for Wordnets co-located with LDK 2017: The first conference on Language, Data and Knowledge, as well as the three anonymous reviewers of the present article.

The authors declare that they have no competing interests.

The authors’ contribution was as follows: concept of the study: Valentina Stefanova (60%), Tsvetana Dimitrova (40%); data analyses: Tsvetana Dimitrova (50%), Valentina Stefanova (50%); the writing: Tsvetana Dimitrova (70%), Valentina Stefanova (30%).

This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 PL License (http://creativecommons.org/licenses/by/3.0/pl/), which permits redistribution, commercial and non-commercial, provided that the article is properly cited.

© The Authors 2018

Publisher: Institute of Slavic Studies, Polish Academy of Sciences, University of Silesia & The Slavic Foundation