Proceedings of the ESSLLI 2022 Student Session, 97–109
doi.org/10.21942/uva.20370402

Towards incorporating prepositions in BTB-WordNet: A case study

Zara Kancheva

Institute of Information and Communication Technologies
Bulgarian Academy of Sciences, 1000 Sofia, Bulgaria
iict@bas.bg
https://www.iict.bas.bg/en/index.html
zara@bultreebank.org

Abstract. The paper presents an initial attempt at investigating the topic of preposition incorporation in BulTreeBank WordNet. The focus is on the preposition 'на' (*na*) in *verb+preposition+noun* constructions. The prepositions are categorised by semantic classification and the verbs by the categories from Princeton WordNet. Wordnets typically do not contain prepositions, because they are difficult for processing, but their incorporation would seriously benefit the performance of wordnets for tasks such as text analysis and generation, word-sense disambiguation, automatic translation, etc.

Keywords: WordNet · Preposition · Word-sense disambiguation

1 Introduction

Prepositions have a substantial role in many natural language processing tasks, but their polysemy constitutes one of the greatest challenges for this research area and a considerable amount of work has been done towards their classification and disambiguation. They have an important role in every language, but in analytic languages such as Bulgarian it is even more considerable – they express the semantic and syntactic relations between words in phrases [2].

These features of the prepositions determine their importance in the field of NLP, but also the challenges related with their processing – they are one of the word classes which are both frequent and extremely ambiguous [6]. Prepositions have significant impact in tasks like word-sense disambiguation, machine translation, syntactic parsing, prepositional phrase attachment resolution, knowledge extraction and word embeddings. PrepNet, originally created for French, seems to be the only pure preposition resource, which could be used for different languages [19].

This word class usually is missing in wordnets, but its incorporation there would be very useful – a good coverage and representation of prepositions in a wordnet would seriously extend its circle of applications and would improve its performance if used for text analysis and generation, word-sense disambiguation, automatic translation, etc. The BulNet[1] for Bulgarian seems to be an exception, but the prepositions in it are presented as they usually are in explanatory dictionaries (definition, synonyms and examples) and do not have any relations with other synsets (neither between the prepositions, nor with other parts of speech). A translation of the corresponding English prepositions is also given. Such a presentation is a good step forward, but for the NLP tasks more information is needed. In addition to the definitions and examples, the preposition synsets in BTB-WN will have relations with other prepositions and with different parts of speech. Also, they will have synset categories based on the sense that they express.

This paper is an initial attempt at investigating the topic of ambiguous preposition incorporation in BulTreeBank WordNet (BTB-WN)[16]. The focus is on one particular preposition and one type of phrase construction – the preposition 'на' (*na*) and constructions in which the preposition is used with verbs. The choice of the phrase type excludes the processing of adjuncts, because they would not benefit the intended verb analysis, since they do not specify the semantics of the verb as opposed to the complements, which are planned to be used for analysis in the future work. 'На' (*na*) is the most frequent and most polysemous Bulgarian preposition and its translation into English depends on the context, but commonly it is translated as 'of', 'on', 'to', which are also highly ambiguous.

I assume that the strategy I propose for the incorporation will be applicable also to other ambiguous prepositions. The first step is the categorisation of the preposition towards a Bulgarian semantic classification and of the verbs from the phrases by wordnet categories. The aim is to explore whether there are verb categories which typically combine with 'на' (*na*) and express particular sense. Then the prepositions will be integrated in BTB-WN with definitions and relations between them and verbs will be established. For this task the available resources for Bulgarian and the connections between them will be used – the BulTreeBank [18] and BTB-WN, and in the future also the OntoValence lexicon [15]. The data is extracted from the BulTreeBank and the BTB-WN is used for the analysis of verb categories. The extraction is performed with the XML-based CLaRK system.[2] The interconnectedness of these resources for Bulgarian is a serious advantage for the task – the syntactic annotations in the treebank (both Universal Dependency and head-driven phrase structure grammar annotations) are connected with the valency lexicon and with BTB-WN on the level of sense.

Section 2 gives an overview of different preposition research, Section 3 presents the semantic preposition classification that is used, Section 4 contains the analysis and results, and Section 6 concludes the paper.

[1] http://dcl.bas.bg/en/resursi/wordnet/
[2] http://bultreebank.org/bg/clark/

2 Related Works

Bulgarian prepositions are very well explored: a thorough review of the history of preposition classifications is presented in [2]. The work of [3] is focused on the prepositions with causal semantics. In [5] a comparative analysis of prepositions in Balkan languages is performed. Frequently the research are dedicated to particular prepositions. Many related works ([10], [12], etc.) propose diverse preposition classifications, which could be applied also to Bulgarian prepositions, but here a specific classification for the preposition 'на' (*na*) will be used.

There are various research related with preposition processing for NLP and preposition incorporation in wordnets. Most of them take into consideration a particular preposition (as the current research) or a preposition category ([9] focuses on the preposition 'of'; [8] on Portuguese prepositions for movement; [6] on English prepositions 'for', 'with', 'at' 'on' and their equivalent Tamil postpositions). Sometimes prepositional phrases with a few elements are excerpted and analysed [9], but whole sentences ([6], [8]) are also observed. Here whole sentences are extracted. Some of them use lexico-semantic resources like VerbNet[3], FrameNet[4], PropBank[5], etc.

One of the fundamental works on preposition application in wordnet is [9] where by using information from Princeton WordNet and applying inferential heuristics two types of phrases are analysed – *noun+preposition+noun* and *verb+preposition+noun* – and are organized in classes if there are hyperonymy, hyponymy or synonymy relations between the verbs and the nouns in the phrases or if they have common hypernym/hyponym. This research considers also the synset definitions, for instance, if the definition starts with 'the act of', it could be concluded that the given noun denotes an action; or if a verb definition contains given nouns it is very likely that they are its objects; and if a verb definition begins with other verb, probably both verbs denote the same action, but in different context. This approach is beneficial with wordnets that have systematically built definitions following certain models of presentation. Such analysis on the base of definitions could not be performed in the current research, because BTB-WN has been built in several stages, it is mapped with different resources, various annotators have worked on its creation and extension, and the construction of uniform definitions has not been strictly followed. However, the approach of [9] towards definitions, could be performed in the work with BTB-WN through its mapping with OEW, if such tendencies are available there.

The work of [8] presents a very interesting approach towards preposition integration in wordnet, including visual description – Portuguese prepositions and multi-word expressions (that function as and could be replaced with simple prepositions) which express movement are incorporated through typical wordnet

[3] https://verbs.colorado.edu/verbnet/
[4] https://framenet.icsi.berkeley.edu/fndrupal/
[5] https://propbank.github.io/

relations: synonymy, antonymy, hyponymy/hyperonymy and causes/is caused by. The relations are established by a series of tests, for example 'under' and 'on top of' are antonyms if 1)'under' and 'on top of' are both hyponyms of 'in', 'at'; and 2) 'under the table' is the opposite of 'on the top of the table' and 'on the top of the table' is the opposite of 'under the table'. Another work on preposition ambiguity resolution with focus on machine translation and involvement of wordnet is [6]. *Hypernym* and *Lexicographer file* information (in the lexicographer files synsets are structured based on syntactic category and logical grouping) is extracted from PWN and used for machine learning models together with dependency and collocation information from other sources.

PrepNet [10] is a repository of preposition syntactic and semantic behaviours. The paper contains a detailed semantic categorization of prepositions, inspired from thematic role classifications, but for the purpose of the research a Bulgarian-specific classification will be used. Senses are organized on three levels: first level characterises semantic families, second level accounts for different facets of the semantic families, and third level the modalities of a facet when appropriate. The classification proposed in PrepNet could be used in subsequent stages of my research, because BTB-WN is mapped with OEW, so such information for English prepositions will be useful.

For the purpose of resolving prepositional ambiguity and improving syntactic parsing [7] obtain selectional restriction information from VerbNet and WordNet. [12] present a semantic classification of prepositions using the semantic roles from Penn Treebank and FrameNet, and additionally they use the high-level synsets of PWN as word classes.

The Preposition project is a noteworthy resource of preposition senses suitable for NLP [14]. The senses there are described by dictionary definition, basic syntactic function and meaning, other prepositions with similar semantic role, data from the FrameNet inventory, different syntactic forms with which the semantic roles may be realized, and position in a network of prepositions.

My research is directed towards wordnet like [9], [8], [11] and [6], but it uses different wordnet feature in the analysis of prepositions – the verb categories. The approach of [11], [12] and [14] for semantic role classification of prepositional phrases with lexical and semantic data from resources such as FrameNet, VerbNet and PropBank will be adopted in the future work on BTB-WN.

3 Classification

In this section I outline the semantic classification of Bulgarian prepositions and the verb categories. My hypothesis is that each of the classes of verbs determines certain semantic properties of the preposition in phrases of the type *verb-preposition-noun*. I explored such combinations extracted from BulTreeBank, identified the relation that the preposition expresses and the verb class.

3.1 Classification of prepositions in Bulgarian

Prepositions in Bulgarian grammar are typically classified by their origin, morphological composition and semantics. The categorization proposed in [1] contains thirteen semantic types: *locative*; *temporal*; *manner*; *causal*; *purpose*; *tool and instrument of action*; *reason, opinion, compliance*; *prepositions for denoting object of thought*; *possessive*; *origin and part of a whole relation*; *quantitative*; *surpassing a limit*. Not all of these classes are applicable to the preposition 'на' (*na*).

In [4] it is stated that 'на' (*na*) is the most frequent Bulgarian preposition (this claim is also supported in the BulTreeBank corpus), because it is loaded with the most numerous and most abstract senses, and presents a very detailed classification of its senses (it covers eight from the above mentioned thirteen sense categories). Several semantic usages of this poly-functional preposition are outlined there, but additionally it includes one of its most frequent functions – the expression of indirect object. Since wordnet is a resource based on semantic relations, the semantic classes of prepositions are more relevant for this work than their syntactic functions and some problematic (ambiguous or not purely semantic) classes from this categorization will be reconsidered. The classification contains the following types:

1. *property and possession* relations (for example, Почеркът е на жена, *Počerkăt e na žena*, 'The handwriting is of a woman'),
2. in combination with nouns for the *expression of indirect object*: (a) after the preposition there are the person or object upon which the action goes or towards which it is directed (Той пишеше на баща си отчаяни писма за пари, *Toj pišeše na bašta si otčajani pisma za pari*, 'He wrote to his father desperate letters for money'), and (b) with some verbs like уча, *uča*, 'teach', навиквам, *navikvam*, 'get used to', мириша, *miriša*, 'smell' the preposition inserts the indirect object which is not object of the action, but an explanation of the actions' features (...уж да ги научат на почит към старите..., *už da gi naučat na počit kăm starite*, 'ostensibly to teach them respect for the elderly'),
3. *locative* relations – (a) place in which limits or surface a given action is performed, and (b) direction of action (Той дълго стоя на шосето и гледа подир каруцата, *Toj dălgo stoja na šoseto i gleda podir karucata*, 'He stood on the road for a long time and watched after the cart'),
4. *purpose* relations – (a) direction of action according to some purpose, (b) predefined goal, role, function of something, and (c) reason for the action (Ако ми дадеш билет, бих дошъл на театър, *Ako mi dadeš bilet, bih došăl na teatăr*, 'If you give me a ticket, I would come to theatre'),
5. *temporal* relations – (a) position in time, period of time during or around which something is performed, direction of action towards given moment, marking of dates and others, (b) simultaneous expression of time and circumstances in which the action is performed, and (c) periodic change of time

parts in combination with preposition ‘от’, *ot* (...сватбата на Катрин Зита-Джоунс и Майкъл Дъглас, която е на 18 ноември, *svatbata na Katrin Zita-Džouns i Majkăl Dăglas, kojato e na 18 noemvri*, ‘Catherine Zeta-Jones and Michael Douglas’s wedding, which is on 18 November’),

6. *compatibility with action* – expression of circumstances in which the action takes place or which are conditions for its occurrence (блясваха снежно бели на месечината, *bljasvaha snežno beli na mesečinata*, ‘glowing snow white on the moon’),[6]
7. *quantitative* relations – (a) indication of age or presentation of age as a period of time in people’s life, (b) measure, degree, quantitative increase, distance size, and (c) disintegration of the whole into parts (‘Справяне с предизвикателствата на живота’ може грубо да се раздели на три части’, *‘Spravjane s predizvikatelstvata na života’ može grubo da se razdeli na tri časti* ‘*Dealing with the challenges of life* can be roughly divided into three parts’),
8. *manner* relation (Понякога ми липсва даже усещането да се чувствам на върха, когато достигам до решението на геометрична задача..., *Ponjakoga mi lipsva daže useštaneto da se čuvstvam na vărha, kogato dostigam do rešenieto na geometrična zadača*, ‘Sometimes I even miss the feeling of being on top when I come to the solution of a geometric problem’),
9. *transition* in new, different state with verbs like превръщам се, *prevrăštam se*, ‘turn in’, ставам, *stavam*, ‘become’, преструвам се, *prestruvam se*, ‘pretend’ (Кажи му да се търкулне в леглото и тъй да се престори на болен, *Kaži mu da se tărkulne v legloto i tăj da se prestori na bolen*, ‘Tell him to roll over in bed and pretend to be sick’),
10. *instrument of action* expression (...никой не умееше да свири на кавал като него, *nikoj ne umeeše da sviri na kaval kato nego*, ‘no one could play the kaval like him’).

All of the *purpose* subtypes will be processed as general purpose relation, because their meaning could be summarised as goal of motion.
The *compatibility with action* sense is one of the rarest and is characteristic of expressive speech. The examples of this class bear location semantics, so they will be included in the location relation.
The *instrument of action* expression category can be viewed as indirect object expression and will not be analysed as a separate class.
Some examples for the *manner* relation are ambiguous, often because they are figuratively used. To some extent they could be interpreted as locative (for example in Голяма част от тези реформи съществуват само на хартия..., *Goljama čast ot tezi reformi săštestvuvat samo na hartija*, ‘Many of these reforms exist only on paper’), but here they will be considered as manner relations.

[6] The example is taken from [1].

3.2 Classification of verbs

The categories from PWN[7] which are inherited in OEW and BTB-WN will be used for the verb analysis. There 15 classes of verbs are presented: *verb.body* (verbs of grooming, dressing and bodily care); *verb.change* (verbs of size, temperature change, intensifying, etc.); *verb.cognition* (verbs of thinking, judging, analyzing, doubting); *verb.communication* (verbs of telling, asking, ordering, singing); *verb. competition* (verbs of fighting, athletic activities); *verb.consumption* (verbs of eating and drinking); *verb.contact* (verbs of touching, hitting, tying, digging); *verb.creation* (verbs of sewing, baking, painting, performing); *verb.emotion* (verbs of feeling); *verb.motion* (verbs of walking, flying, swimming); *verb.perception* (verbs of seeing, hearing, feeling); *verb.possession* (verbs of buying, selling, owning); *verb.social* (verbs of political and social activities and events); *verb.stative* (verbs of being, having, spatial relations); and *verb.weather* (verbs of raining, snowing, thawing, thundering).

4 Analysis and Results

750 sentences with phrases of the type *verb*+na+*noun* were extracted from BulTreeBank. Some of them contain more than one phrase of this type, so the final number is 768. The first 210 of them were manually sorted following the classification of [4]. This part of the data contains fiction, legal documents, popular science and journalistic texts. The annotation was performed by the author of the paper and thus the initial problematic cases were successfully identified. However, the rest of the data is planned to be validated by more people. The results show that most usages of 'на' (*na*) are from the category for *expressing indirect object*, followed by the *locative* and *purpose* categories. Several categories did not find examples in the used data and their absence leads to some conclusions.

On one hand, the nature of the data has a high impact on the results – some of the senses of 'на' (*na*) are characteristic mainly for fiction (*compatibility with action*), others for archaic texts (*predefined goal*, *role*, *function* of something), so they would have more occurrences in specific corpora. Some senses from the classification could be expressed only with particular verbs (*transition* in new, different state; *indirect object expression* with verbs like уча, *uča*, 'teach', навиквам, *navikvam*, 'get used to', мириша, *miriša*, 'smell', etc.).
On the other hand, the restriction of the phrase type also influences the results – some senses of the preposition could not be expressed or rarely are when it is used with verbs. For instance the *property and possession* sense of 'на' (*na*) is a very productive class, but in the analysed data only three examples are found, so it could be assumed that *property and possession* relations are more frequently expressed with nouns. Of course, this is not the case with every category. The

[7] https://wordnet.princeton.edu/documentation/lexnames5wn

temporal, *quantitative* and *manner* senses are generally productive, but in the current data they have only a few examples.

There are only six examples of 'на' (*na*) expressing *property and possession* relations – three of them are with the verb 'be', which is categorised as *verb.stative*.

The *expression of indirect object* with 'на' (*na*) is not of interest to the research, but still the results could be summarised – this class has 120 examples with the greatest variety of verbs, most of them from *verb.possession*, *verb.communication* and *verb.stative* categories.

The *locative* examples (40) follow a clear pattern – most of the verbs there are from the *verb.contact* (hang, lean, stay, lie, put, bang, etc.) and *verb.motion* (go, stay, sit, return, go out, land, etc.) categories, and a few from *verb.stative* (live, attend, be). There is a relevantly big group of phrases with the verb 'be', 'на' (*na*) and a spatial noun (for instance, Стаята ни беше на горния етаж, *Stajata ni beše na gornija etaž*, 'Our room was on the upper floor'). Three examples were found for the more specific locative relation – *direction of action* – and they contain the verbs ида, *ida*, 'go', избягвам, *izbjagvam*, 'run away' (both *verb.motion*) and оглеждам се, *ogleždam se*, 'look around' (*verb.perception*).

Purpose relations (18 examples) are mainly expressed with *motion* and *social* verbs, the verbs отивам, *otivam*, 'go' and съм, *săm*, 'be' with nouns such as работа, *rabota*, 'work', събрание, *săbranie*, 'meeting' and заседание, *zasedanie*, 'sitting' which could not be interpreted as locations, but rather as purposes – отивам на работа, *otivam na rabota*, 'I am going to work (noun)' should be interpreted as отивам да работя, *otivam da rabotja*, 'I am going to work (verb)'.

The semantics of the verbs in the *temporal* category is not surprising – започвам, *započvam*, 'start' (*verb.change*), съм, *săm*, 'be' and свършвам, *svăršvam*, 'end' (*verb.stative*) – but the number of examples in the data is rather unexpected. *Temporal* relations for position in time and marking of dates ('на' (*na*) is the only preposition that introduces dates) expressed by 'на' (*na*) are generally very frequent, but only two examples were found. The other two *temporal* subtypes are rare and do not have any examples in the analysed data – first of them could be realised only with verbal nouns and the second – only in given phrases in combination with the preposition 'от' ('from').

Quantitative relations (12 examples) are frequently expressed with 'на' (*na*). It is the only preposition for expression of age in combination with the verb съм, *săm*, 'be' (*verb. stative*); in the data there are two examples of this kind.
The subtype for *measure, degree, distance*, etc. is realized with the verbs съм, 'be', ставам, *stavam*, 'happen' and нагласям, *naglasjam*, 'set' (both *verb.change*). The most productive quantitative sense in the data is for *disintegration of the whole into parts* and it is expressed with semantically related verbs – разцепвам се, *razcepvam se*, 'split up', деля се, *delja se*, 'divide' (both *verb.social*), разпадам се, *razpadam se*, 'disintegrate' (*verb.change*), разделям, *razdeljam*, 'separate' (*verb.cognition*).

Table 1. Verb classes distribution in the preposition categories and number of phrase occurrences

Prepositional category	**Verb categories**	**Number of occurrences**
locative	*verb.contact* *verb.motion* *verb.stative* *verb.perception* *verb.body* *verb.creation*	40
purpose	*verb.stative* *verb.motion* *verb.social* *verb.change*	18
quantitative	*verb.stative* *verb.change* *verb.social* *verb.cognition*	12
property and possession	*verb.change* *verb.stative* *verb.consumption* *verb.cognition*	6
manner	*verb.cognition* *verb.emotion* *verb.creation* *verb.stative*	4
transition	*verb.communication*	3
temporal	*verb.stative* *verb.change*	2

Manner relation (4 examples) is expressed with verbs from various categories – оставям, *ostavjam*, 'leave' (*verb.cognition*), чувствам, *čuvstvam*, 'feel' (*verb.cognition*), извършвам, *izvăršvam*, 'perform' (*verb.creation*) and съществувам, *săštestvuvam*, 'exist' (*verb.stative*).

The examples for *transition* in new state relation are three and they contain the same verbs mentioned in [4] and prove that the class has a narrow circle of compatible verbs – правя се, *pravja se*, направя се, *napravja se*, преструвам се, *prestruvam se*, 'pretend' (*verb.communication*).

On the base of this analysis the following features are used to identify the preposition sense and represent it in BTB-WN. For instance, from the example sentence

(1) Kolednata vakancija šte započne na 21 dekemvri i
Christmas-ADJ-DEF vacation start-FUT-SG on 21 December and
šte svărši na 14 januari
end-FUT-SG on 14 January
'Christmas vacation will start on 21 December and end on 14 January.'

we can observe that the verbs from the phrase with 'на' (*na*) bear temporal semantics, they are categorised as *verb.change* and *verb.stative* in BTB-WN. The noun phrases are also related with time – they contain numerals and nouns for months. Numerals are rarely included in wordnets (mainly cardinal numbers), but for the nouns there is a category *noun.time* which could be used for such cases. From the semantics of the verbs and nouns in the example it could be considered that the use of 'на' (*na*) expresses temporal relation. For this meaning *na* will have a synset with definition 'Indication of the time when something is happening, is being done' and *temporal* relation with synsets of verbs like 'start' and 'end'.

An example could be made also for the *locative* relation:

(2) Tja zastana smajana na praga
3SG-F stand-PST-SG stunned-F-SG on doorway-DEF
'She stood stunned on the doorway.'

In wordnet заставам, *zastavam*, 'stand' is *verb.contact* and the noun праг, *prag*, 'doorway' is *noun.artifact*, but its spatial semantics is easily observed. For this sense of *na* the synset will have a definition 'Indication of the place where something is happening or where someone or something is located' and *locative* relation with *contact* verbs such as 'hang', 'stand', 'sit', 'lie'.

These two examples show that my hypothesis is feasible. I will apply this strategy to all verb classes for which it is applicable. For the rest of the *verb-proposition-noun* I assume that I will need more detailed classification of the semantic roles of the corresponding verbs. Here I will rely on the valency lexicon. Where necessary I will perform manual classification. Following this strategy I hope to minimize the manual work in the task of preposition incorporation within BTB-WN.

5 Towards a model of the preposition synsets

The prepositions in BTB-WN will be presented with a detailed definition of the relation they express, synonyms if available, and examples. A generalised classification which is applicable for all Bulgarian prepositions will be created and the types will be used as categories of the synset. Two kinds of synset relations will be considered: relations between preposition synsets, and between prepositions and other parts of speech (in the first place between verbs and prepositions which together express a given meaning, such as the above mentioned verbs 'pretend' and 'turn in' which combined with на express transition in new, different state). For example the synset of на ('on') with definition "Location relation in which something is located on some surface" has a synonym – the preposition върху, *vărhu*, 'on', and also has an antonym relation with the synset for под, *pod*, 'under'.

6 Conclusion

The current work serves as an initial step towards the large-scale integration of prepositions in BTB-WN. The approach combines a semantic preposition classification, verb categorisation from wordnet and in the next stage also the relations and noun categories from wordnet, and features from a valency lexicon will be taken into consideration.

The analysis shows that there are some distinguishable tendencies – particular groups of verbs tend to combine with the preposition 'на' (*na*) for specific senses, so these relations will be further explored and the analysis will be elaborated with information from the hierarchy inheritance of the verbs and nouns in wordnet and the valency lexicon. Since BTB-WN is mapped with the Open English WordNet the prepositions would be considered in a cross-lingual aspect as part of future work.

Acknowledgements This work was supported by *the Bulgarian National Interdisciplinary Research e-Infrastructure for Resources and Technologies in favor of the Bulgarian Language and Cultural Heritage, part of the EU infrastructures CLARIN and DARIAH – CLaDA-BG*, Grant number DO01-377/18.12.2020. and the Grant No. BG05M2OP001-1.001-0003, financed by the Science and Education for Smart Growth Operational Program (2014-2020) and co-financed by the European Union through the European structural and Investment funds.

References

1. Boyadzhiev, T., Kutsarov, I., Penchev, Y.: Contemporary Bulgarian language. Phonetics, lexicology, word formation, morphology, syntax. Petar Beron, Sofia (1998) (in Bulgarian)
2. Konstantinova, V.: Prepositions in Bulgarian grammar literature. Publishing house of BAS, Sofia (1982) (in Bulgarian)
3. Konstantinova, V.: Prepositions with causal function in contemporary Bulgarian standard language. Publishing house of BAS, Sofia (1983) (in Bulgarian)
4. Stoyanov, S. Ed.: Grammar of contemporary Bulgarian standard language. Morphology. Publishing house of BAS, Sofia (1983) (in Bulgarian)
5. Asenova, P.: Community in the use of the most characteristic grammaticalised prepositions in Balkan languages. Yearbook of Sofia University, Faculty of Slavic studies, vol. 66/1, pp. 179-250 (1972) (in Bulgarian)

6. Kumar, M., Rajendran, S., Soman, K. P.: Cross-Lingual Preposition Disambiguation for Machine Translation. Procedia Computer Science, vol. 54, pp 291-300. (2015) 10.1016/j.procs.2015.06.034.
7. Bailey, D., Lierler, Y., Susman, B.: Prepositional Phrase Attachment Problem Revisited: how Verbnet can Help. In Purver, M., Sadrzadeh, M., Stone, M. (eds.) Proceedings of the 11th International Conference on Computational Semantics, pp. 12–22. Association for Computational Linguistics, London, UK (2015)
8. Amaro, R.: Integrating Prepositions in Wordnets: Relations, Glosses and Visual Description. In: Calzolari, N., Choukri, K., Cieri, C., Declerck, T., Goggi, S., Hasida, K., Isahara, H., Maegaard, B., Mariani, J., Mazo, H., Moreno, A., Odijk, J., Piperidis, S., Tokunaga, T. (eds.) Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). European Language Resources Association (ELRA), Miyazaki, Japan. (2018)
9. Harabagiu, S. M.: An application of WordNet to prepositional attachment. In Proceedings of the 34th annual meeting on Association for Computational Linguistics (ACL 1996), pp. 360–362. Association for Computational Linguistics, Santa Cruz, USA. (1996) https://doi.org/10.3115/981863.981915
10. Saint-Dizier, P.: PrepNet: a Framework for De-scribing Prepositions: Preliminary Investigation Results. In Fifth International Workshop on Computational Semantics (IWCS05), Tilburg (2005)
11. O'Hara, T., Wiebe, J.: Exploiting semantic role resources for preposition disambiguation. Computational Linguistics, vol. 35(2), pp. 151–184 (2009)
12. O'Hara, T., Wiebe, J.: Preposition semantic classification via Penn Treebank and FrameNet. In: Proceedings of the Seventh conference on Natural language learning at HLT-NAACL 2003 - Volume 4 (CONLL '03), pp. 79-86. Association for Computational Linguistics, USA. (2003) DOI:https://doi.org/10.3115/1119176.1119187
13. Baldwin, T., Kordoni, V., Villavicencio, A.: Prepositions in Applications: A Survey and Introduction to the Special Issue. Computational Linguistics, vol. 35, pp. 119-149. (2009)
14. Litkowski, K. C., Hargraves, O.: The Preposition Project. In ACL-SIGSEM Workshop on "The Linguistic Dimensions of Prepositions and Their Use in Computational Linguistic Formalisms and Applications", University of Essex - Colchester, United Kingdom, pp. 171-179. (2005)
15. Osenova, P., Simov, K., Laskova, L., Kancheva, S.: A Treebank-driven Creation of an OntoValence Verb lexicon for Bulgarian. In: Calzolari, N., Choukri, K., Declerck, T., Doğan, M. U., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S. (eds.) Proceedings of LREC'12, pp. 2636-2640. Istanbul, Turkey, ELRA (2012). 978-2-9517408-7-7
16. Osenova, P., Simov, K.: The data-driven Bulgarian WordNet: BTBWN. Cognitive Studies – Études cognitives, 2018(18) (2018) https://doi.org/10.11649/cs.1713
17. Cassidy, P. J.: An investigation of the semantic relations in the Roget's Thesaurus: Preliminary results. In Proceedings of the First International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2000), pp. 181–204. Mexico City (2000)
18. Simov, K., Popova, G., Osenova, P.: HPSG-based syntactic treebank of Bulgarian (BulTreeBank). In: Wilson, A., Rayson, P., McEnery, T. (eds.) A Rainbow of Corpora: Corpus Linguistics and the Languages of the World, pp. 135-142. Lincom-Europa, Munich (2002)

19. Saint-Dizier, P.: Syntactic and Semantic Frames in PrepNet. In: Proceedings of the Third International Joint Conference on Natural Language Processing: Volume-II, pp. 763-768. Hyderabad, India (2008)