

CORPUS-BASED SEMANTIC MODELS OF THE NOUN PHRASES CONTAINING WORDS WITH 'PERSON' MARKER

MARGARYTA LANGENBAKH

Taras Shevchenko National University of Kyiv, Ukraine

LANGENBAKH, Margaryta: Corpus-Based Semantic Models of the Noun Phrases Containing Words with 'Person' Marker. *Journal of Linguistics*, 2017, Vol. 68, No 2, pp. 249 – 257.

Abstract: The mechanism underlying constructing of lexically correct sequences of words is an object of attention both in theoretical and applied fields of linguistics. This paper reveals some aspects of modelling the patterns of semantic valence in noun phrases of NN (Noun+Noun) structure, one or both components of which contain the 'person' semantic tag. The research is based on the Corpus of Ukrainian and performed with the help of automatic language processing.

Keywords: natural language processing, semantic valence, noun phrases

1 COMPUTER-BASED REPRESENTATIONS OF SEMANTICS

The problems we discuss lie at the intersection of important theoretical (syntactical and lexical valence theory) and practical (semantic and informational analysis, text mining) studies. Last three decades in the computational linguistics brought a considerable amount of studies and projects representing different approaches to the meaning representation: the WordNet [17], FrameNet [3], Wikipedia-based annotation [20], Abstract Meaning Representation (AMR) [12], the method of propositions [20] etc. The WordNet project initiated the creation of numerous set of semantic dictionaries all over the world including Slavic languages that are represented by the Polish [8], Slovene [9], Croatian [19], Bulgarian [13] or Russian [4] projects. During the last decade some researches have been taken on the base of Ukrainian language [1], [14].

Active development of the corpus linguistics actualized a new field of study – the semantic annotation of the textual corpora. Now we have a semantic annotation for the Bulgarian [13], Polish [5], Russian [15] and Slovene [9] corpora. The semantic annotation for the Corpus of Ukrainian is being developed by the National Taras Shevchenko University of Kyiv [6]. In the next sections we discuss some theoretical and practical aspects of the taxonomy-based approach to the semantic annotation of the corpus by the example of noun phrases modelling.

2 THE THEORETICAL ASPECTS OF RESEARCH

Traditionally computer-based semantic analysis (as long as syntactic) is performed at the sentence level. But choosing the phrase as a unit of semantic processing has some advantages. Assigning the semantic roles to the sentence constituents is not a problem when we deal with monopropositional constructions. Polypropositional

sentences may contain a set of linked situations and the word included in more than one situation may have different roles in them. For example, in the sentence *Я не вважаю твого брата винуватцем сварки* 'I don't think that your brother caused a quarrel', the noun *брата* 'brother' is an object of one situation (to think him not to be a causer) and a subject of another (to cause a quarrel). Using a phrase as basic semantic construction may solve this problem.

The goal of our research was to investigate the mechanisms of semantic connection inside the noun phrases structure. In order to do this, we had to answer several questions.

The first question what taxonomic markers are needed to form the semantic models of noun phrases and if some fixed positions (head or adjunct).

Every natural language has some syntactical rules. In our article we tried to find out whether these rules are based on the grammar only or there are some lexical features that determine the words positions.

The second question was what combinations of the semantic tags correlate with the certain types of relations.

One more important question was whether the core meaning markers determine the relation or peripheral elements can also take part in this process. Many researchers of the lexical system emphasize on the different roles of semantic components of the word meaning. According to Y. Apresyan, "every lexical meaning has certain syntactical structure" [2, p. 9]. Examining the structure of lexeme, the linguists divide it into integral and distinctive semantic features [21, p. 78]. The first of them help to specify the semantic similarity between the words and the second – to distinguish the meanings of semantically similar lexemes. And one of our tasks was to find out if there is a correlation between the position of the semantic marker in the meaning structure and the word valence.

The last question we had to answer was what are the functions of the semantic model elements. E. Sapir proposed to differentiate the meanings into concrete (nominative, descriptive), derivational, relational and mixed concrete relational [11, p. 20]. The presence of relational markers in the lexemes gives them a potentiality to attach the other lexemes (the valence phenomenon [23], [13]). The functional structuring of the word semantics also plays an important role in the development of ontologies. For example, in the WordNet project the elements that have relational meanings are used to represent the semantic relations between the nominative elements [17]. In our research we tried to find out which of the meaning components determine the nature of syntagmatic links – relational that initiate the connection or nominative that define the nature of word lexical meaning.

3 TYPES OF RELATIONS AND THEIR SEMANTIC MARKERS IN THE NN PHRASES CONTAINING WORDS WITH 'PERSON' TAG

In our previous work we analysed the deep syntactic structure patterns of the noun phrases [16]. This research is a second attempt to investigate the mechanisms of the noun valence realization focused more on the lexical background of this phenomenon and performed with the help of machine semantic tagging.

We used the taxonomy-based semantic dictionary for automatic semantic analysis of the texts. The taxonomy consists of three general classes (Proper Names, General Names and Object Names), divided into a list of subclasses (178 taxons, the maximum depth of taxonomy – four levels). The dictionary was applied to the newspapers subcorpus (17 705 122 words) of the Corpus of Ukrainian (<http://mova.info/corpus.aspx>) processed by the AGAT NLP-system [7]. Then we automatically selected the phrases of the NN structure with their grammatical and semantic tags (Fig. 1). The query to the database returned 113 471 phrases consisted of 12 610 words¹.

Paryslovpredl_nt										
text_id	connect_ty1	word	lemm	word2	code	semcode1	lemm2	code2	semcode2	
5996 ICL		частині	частина	Термінатора	КП	1) class,pt0part 2) class,pt	термінатор	ЙР	51) t0space 51	
5996 ППЗ		до	до	ієрархії	ЛЩПР	51) t0sound	ієрархія	КР	51) t0hum,pt0c	
5996 IC36		символ	символ	партії	ЙИ	1) t0text0figure 54) t0mer	партія	КР	2) pt0aggr 4) ф	
5996 IC-pr		прапора	прапор	піратів	ЙР	1) t0text0figure 2) t0ment	пірат	ЙЕ	1) t0hum	
5996 DCLS		білі	білі	комірці	ІА	51) t0disease	комірець	ЙА	51) t0toool0dot	
5996 IC38		оформлення	оформлення	коронації	ЛИ	51) der0v,t0change,t0act	коронація	КР	51) t0action	
5996 IC37		систему	система	символів	КВ	4) r0abstr0mereol0set,pt0	символ	ЙЕ	1) t0text0figur	
5996 IC39		промови	промова	пафосу	КУ	1) t0speech 2) t0speech	пафос	ЙР	1) t0condit,t0a	
5996 IC-pr		Фронту	фронт	змін	ЙР	1) r0abstr0mereol0part 3)	зміна	КЕ	51) t0change	
5996 IC-pr		експортера	експортер	електроенер	ЙР	1) t0hum,d0nag	електроенерг	КР	51) r0energy	
5996 IC36		керівник	керівник	експортера	ЙИ	51) t0hum,d0nag	експортер	ЙР	1) t0hum,d0na	
5996 IC-pr		імені	ім'я	партії	ЛР	51) r0propn,t0speech 52)	партія	КР	2) pt0aggr 4) ф	
5996 IC37		діяльністю	діяльність	партії	КТ	1) t0activity 2) t0activity	партія	КР	2) pt0aggr 4) ф	
5996 IC39		майно	майно	партії	ЛТ	51) pt0aggr,t0poss	партія	КР	2) pt0aggr 4) ф	
5996 ICL		Набір	набір	його	ЙИ	53) pt0aggr 111) pt0aggr	його	ККМБМРП	51) t0method	
5996 ICL		Набір	набір	повноважень	ЙИ	53) pt0aggr 111) pt0aggr	повноваженн	ЛЕ	51) r0law	
5996 IC-pr		органів	орган	партії	ЙЕ	1) t0inst 2) t0text 53) t0te	партія	КР	2) pt0aggr 4) ф	
5996 IC-pr		Фронту	фронт	змін	ЙР	1) r0abstr0mereol0part 3)	зміна	КЕ	51) t0change	
5996 IC36		статут	статут	Фронту	ЙИ	1) r0standar	фронт	ЙР	1) r0abstr0mer	
5996 IC36		курс	курс	партії	ЙИ	51) t0move 52) t0idea,r0c	партія	КР	2) pt0aggr 4) ф	
5996 ICL		рішення	рішення	органів	ЛВ	1) der0v,t0event,t0text 2)	орган	ЙЕ	1) t0inst 2) t0t	

Fig. 1. Database structure

According to the results of our research, the relations in noun phrases containing words with the ‘person’ markers may be divided into three general types – actional, non-actional (different sorts of attribution) and part-whole relation (mereology). The actional relations describe the roles of participants in a situation frame. Such structures are the grammatical transforms of sentences (compare *виступ політика* ‘the speech of a politician’ and *Політик виступив* ‘The politician made a speech’). The non-actional relations form the models which in some realizations are similar to the other type of noun phrases – constructions of the noun and adjective (*ввічливість працівника* ‘the politeness of the worker’ is comparable with the *ввічливий працівник* ‘the polite worker’). But this type isn’t completely identical to the noun-adjective attributive relations because some of non-actional phrases also may be transformed into predicative constructions, for example, in a case of possession (*володіння бізнесмена* ‘the possessions of a businessman’ – *бізнесмен володіє (чимось)* ‘the businessman possess (something)’). These examples do not have such definitely actional semantics as the first general type of phrases we described previously, but the attribution characteristic in them has some specificity.

The part-whole relations embody the linguistic representation of mereology phenomenon. They are quite natural for the noun phrases because the ‘part’ and ‘whole’ concepts are usually associated with the objects and described by nouns.

¹ Because of the possible errors of language processing the real results may differ. The estimated error rate is 6–8%

All discussed types of relations can be specified by the semantic patterns they can be used in.

The actional relations are represented by such subtypes:

1. Subjective. These relations appear in the NN phrases that describe the connection between the action and its doer. The actor can possess any grammatical position in the phrase:

A) The position of the subjective grammatical head occurs in such patterns as <creation, person, nomen agentis + action> (*натхненник повстання 'the inspirer of the revolt'*);

<person, nomen agentis + action> (*виконавець проекту 'the executor of project'*) etc. The active mode of the person's role is usually marked by the taxon 'nomen agentis' which describes the doer of certain action;

B) The position of the grammatical adjunct can be illustrated by such pattern as <interaction + person> (*поєдинок чемпіонів 'the battle of champions'*).

Unlike the previous pattern, the adjunct does not obligatory have the 'nomen agentis' marker that can point at the lexical differences between the actions described by these two models.

2. Objective. This subtype is represented by the combination of a word with actional semantics and word with 'person' marker being a passive participant of the situation.

The grammatical position of this element also may be different:

A) Head:

<person + action> pattern (*жертва насильства 'the victim of the violence'*).

B) Adjunct:

<occupation, activity + person> (*лікування пацієнта 'a treatment of the patient'*).

The difference between the lexical meanings of the words that name the actions in subjective and objective models can also produce a semantic variety of actional relations:

– process of creation: <action|process, creation + person> (*підготовка спеціалістів 'a training of the professionals'*);

– contact, interaction: <action|process, contact + person, nomen agentis> (*учасник дебатів 'the participant of debates'*);

– action or process which results in some changes: <action|process, change + person> (*навчання дітей 'teaching of children'*) etc.

3. Subject-object interaction. The noun phrases of this subtype contain the subjective noun that expresses the actor and, at the same time, points to the action, and a noun in an objective role. The 'person' marker can be found in the subjective, objective or both phrase components and the positions of the elements also may vary. The semantic variations are quite wide:

A) Creation of concrete or abstract objects:

<person, nomen agentis, process, creation + tool, furniture> (*виробник меблів 'the manufacturer of furniture'*).

B) Transformation of an object, changing of its characteristics:

<person, nomen agentis, action|process, change + text> (*перекладач роману 'the translator of the novel'*).

- C) Supplying or consumption of products or services:
 <person, nomen agentis + food> (*продавець картоплі* 'a potato seller') etc.
 Our taxonomy doesn't have special tag for naming such class of activity, but these relations are quite regular and semantically different from the others, so we decided to mention them and suppose to add them to the new revised edition of taxons list.
- D) Professional activity:
 <person, nomen agentis, occupation + tool, transport> (*водій тролейбуса* 'a trolleybus driver').
- E) Emotional or willing activity:
 <person, nomen agentis, action|process, will + human qualities|mental sphere> (*захисник моралі* 'the defender of morality').
- F) Participation or collaboration:
 <person, nomen agentis, activity + event> (*учасник фестивалю* 'a participant of the festival').
- G) Professional interaction:
 <person, nomen agentis, occupation + person> (*радник директора* 'the counsellor of the chief manager').

In this subtype (as well as in the other actional constructions) the main lexical meaning of the head is a basis of relational subtype division. For example, in phrases which describe the professional activities the head-actor determines the sense of situation (*приборкувач* – a person who tame an animal) and the adjunct is a tool of concretization and bearer of passive valence marker (*змія* 'snake' is an animal, so it can be an object of taming). The difference between the phrases *творець картини* 'the creator of the picture' and *покупець картини* 'a buyer of the picture' lies in the field of the head's semantics: in the first case we have a person-creator and in the second – a person-consumer. This lexical difference points to the specific mode of the influence which the actor has on the object. The lexical meanings of the adjuncts, on the contrary, don't play an essential role in this type of relations.

The non-actional relations usually are formed by the combination of personal noun and noun describing some characteristics. As well as in previous models, the components of non-actional phrases may appear in different positions: characteristic as the head and subject as adjunct (*самолюбство колег* 'the ambition of colleagues') or inverted variant (*людина честі* 'a man of honour'), but the first pattern is more typical for Ukrainian.

This type of relations also integrates some different semantic variants – complementary, possessive and property attribution.

The property attribution can be found in the noun phrases which contain the word describing specific value of the certain property:

- A) A feature of a human personality:
 <human quality + person> (*велич постаті* 'a greatness of figure').
- B) An emotional state:
 <feeling + person> (*радість матері* 'a joy of mother').

The combination of subjective noun and abstract names of human qualities or activities (fields of knowledge or culture) gives us the complementary attribution (*доктор філософії* 'the philosophy doctor').

The similar situation we have in the case of possessive attributive semantics: *володар зброї* 'the weapon holder'. The possessive relations also can be divided into two subtypes:

A) Ownership:

<person, possession + tool, transport> (*власник автомобіля* 'the car owner').

B) Occupancy:

<place, part of building + person, nomen agentis> (*кабінет секретаря* 'a secretary's room').

As well as the subject-object interaction, the possessive relation of ownership is formed by the subjective noun that has possessive marker. The occupancy case isn't identified by the specific tag of our taxonomy. The choice of the adjunct doesn't rely on its specific lexical meaning but demands a specific relational potentiality (to be an object of possession).

The part-whole relations occur in the patterns, consisting of tags combinations in both phrase constituents. According to the set of markers we divided the patterns in two subtypes:

A) Body and its parts:

<part of the body + person> (*плече робітника* 'the shoulder of the worker'); this pattern includes the tag 'part' of the head and the tag 'person' of the adjunct.

B) Group of people or organization and its members:

<person + person, set of objects> (*член команди* 'a team member'); the head of this pattern has the 'person' taxon (it is not the only semantic component required for such type of relation, but our taxonomy doesn't have a tag for the meaning 'to be a representative of certain group'), the adjunct must be marked as 'set of objects' or 'organization'.

Though the positions of relational markers bearers mainly are not strongly restricted, there are some exceptions when relational type clearly depends on the position of certain marker. For example, a model, built from the head representing the abstract name of class with the 'part' taxon and adjunct with the 'person' and 'set of objects' taxons expresses the 'part-whole' relation (*половина електорату* 'a half of electorate'). At the same time, the converse position of the 'set of objects' marker points to the complementary semantics of relation (*клас професіоналів* 'a class of professionals'). The 'nomen agentis' tag in the phrase head marks its active mode, identifies it as an actor whereas the main elements without this tag are more likely to be an object of an action or a bearer of some qualities. But there are rather rare, minor cases.

There are some situations in which we can't clearly define the type of relation. The first situation is caused by the grammatical homonymy. When the position of a head is occupied by the noun derived from the transitive verb and the adjunct is personal noun we usually are not able to say whether it is an object or subject of an action (*перевірка спеціаліста* 'an examination of the professional' may mean a work of specialist or an examination of his qualification level).

The second situation may be illustrated by the noun phrases, describing the personal relations (*друг сім'ї* 'a friend of family'). The interpretation of such patterns

depends on the specific lexical meaning of the head: in some cases this combination forms the subjective relation (*голова родини* 'a head of family'), in other – objective (*ставленник президента* 'the President's protégé') or even syncretic types (*друг сім* 'i is a possessive model (a friend of whom?), and, from the other hand, subjective, that describes a pattern of behaviour).

Unfortunately, in this article we can't present a full list of the patterns that we found but all of them fit into the three mentioned classes of relations. To sum up our surveillance we can make several conclusions. Firstly, the semantic relation in the noun phrases can be based on the markers of both words (e.g. subjective marker in the head and objective in adjunct) or only one of them. The duplication of certain marker in the phrase constituents rather means their connection abilities than forms the relation (except the case of semantic recursion such as *діти дітей* 'children of children').

Secondly, there are no strict rules concerning the quantity and positions of the markers that determine the relation. In some models they belong to the head, in some to the adjunct or even to both of them. We can only make an assumption that relation is more often based on the element that has stronger valence potential in its meaning (what agrees with the traditional valence theories). Usually they are the names of abstract classes, actors and attributes. Concerning the direction of semantic relation we can draw an analogy with syntax dependencies, divided into three classes – bilateral, unilateral and coordinate [22, p. 89]. The actional relations are usually bilateral because the actor and object of an action are obligatory participants of the situation. The non-actional relations may either be unilateral (in complementary phrases like *категорія читачів* 'a readers category' the head *категорія* 'category' defines the specific role of adjunct *читачів* 'readers') or coordinate (in the attributive phrases like *чесність політика* 'the honesty of a politician' we can say that the grammatical head *чесність* 'honesty' creates the complementary relation with the adjunct *політика* 'politician' and the adjunct also induces a reverse relation of attribution). The part-whole relations are unilateral because the noun which names the 'whole' component of pattern is usually semantically independent.

Thirdly, there is no obvious correlation between the position of the semantic marker in the lexical meaning structure and its potency to establish a relation. In the other words, we cannot say that only heads or adjuncts regularly induce relations. The choice of the semantic markers (and a word correspondingly) needed to build a phrase is determined by the contextual and communicative requirements.

4 THE PROSPECTS OF USING THE TAXONOMY-BASED DICTIONARY IN THE NATURAL LANGUAGE PROCESSING

In the Section 3 we discussed the theoretical issues concerning the dictionary-based valence modelling in the noun phrases. But is the structure of our taxonomic dictionary suitable for automatic detecting the essential lexical features of the words in a context?

Grammatical and semantic annotation theoretically allows the computer to detect the connected words and build the structure of phrases and sentences. Also the

machine can suggest the senses of these constructions according to the markers of their components and relational patterns. The semantic dictionary used in our research embodies the facet classification made by the example of the Russian National Corpora [15], that allows words to be included in several semantic classes and, consequently, to have different sets of markers [6]. It means that we can, treating words as structured semantic complexes, derive the relations between the words not from their integrated lexical meanings but from the certain elements of their semantics that actually provide the semantic connection.

Such approach has some deficiencies. Fixation of all variants of the homonymic and polysemic lexemes leads to high level of the ambiguity and so called informational noise produced by the minor or rare meanings. For example, the word *змiна* 'change' is interpreted by the computer as the 'person' because of its second meaning 'people working in shifts' and, consequently, the phrase *активiсти змін* 'the activists of changes' receives the model <person + person, set of objects> and may be treated as part-whole pattern. Regular wrong results are produced by the connotations of animal names: *використання собак* 'the using of dogs', *кiлькiсть свиней* 'the quantity of pigs' (because in Ukrainian we may use 'a dog' and 'a pig' as abusive words) etc. A majority of words (about 2/3 of dictionary list) has more than one meaning so it is the serious problem.

By applying relational patterns to the processed texts we can reduce the ambiguity to the certain extent. However, there are situations that need more complicated approach – examination of the style, genre, thematic of the text, contextual, statistic or stochastic information etc. Unfortunately, if the ambiguity level is very high such strategy may make the system too complicated.

There are two additional ways to reduce the ambiguity. The first one is to shorten the lexemes description omitting rarely used meanings and very peripheral taxons. The second way is to rebuild the structure of a dictionary. The automatic dictionary should be based not on the full descriptions of the lexical meanings but rather on semantic features that describe words valence. In other words, the definition may be incomplete from the lexical point and include some relational characteristics which lie beyond the lexical meaning. For example, in objective phrase like *споживач пшеници* 'the consumer of wheat' the adjunct must have such marker as 'to be an object of consumption' and its peculiar lexical description doesn't affect the relation (compare: *споживач хлiба* 'the consumer of bread', *споживач енергiї* 'the consumer of energy' etc.). So the machine-oriented dictionary must be more functional and relational than encyclopedical.

References

- [1] Anisimov, A., Marchenko, O., Nikonenko, A., Porkhun, E., and Taranukha, V. (2013). Ukrainian WordNet: Creation and Filling. In Larsen, H. L., Martin-Bautista, M. J., Vila, M. A., Andreasen, T., and Christiansen, H., editors, *Flexible Query Answering Systems. FQAS 2013. Lecture Notes in Computer Science, vol 8132*, pages 649–660, Springer, Berlin – Heidelberg, Germany.
- [2] Apresyan, Yu. (2014). *Izbrannye trudy, tom I. Leksicheskaya semantika. Sinonimicheskie sredstva yazyka*. Izdatelskaya firma «Vostochnaya literatura» RAN, Moscow.

- [3] Baker Collin, F., Fillmore, Ch. J., and Lowe, J. B. (1998). The Berkeley Framenet project. In *COLING-ACL '98: Proceedings of the Conference*, pages 86–90, Association for Computational Linguistics, Montreal, Canada.
- [4] Balkova, V., Sukhonogov, A., and Yablonsky S. (2004). Russian wordnet: From UML-notation to Internet/Intranet Database Implementation. In *Proceedings of the Second International WordNet Conference*, pages 31–38, GWC, Brno, Czech Republic.
- [5] Broda, B., Piasecki M., and Maziarz, M. (2010). Evaluating LexCSD – a weakly-supervised method on improved semantically annotated corpus in a large scale experiment. In *Proceedings of Intelligent Information Systems (2010)*, pages 63–76, Publishing House of University of Podlasie, Siedlce, Poland.
- [6] Darchuk, N., Zuban', O., Lanhenbakh, M., and Khodakiv's'ka, Ya. (2016). AGAT semantyka: semantyczne rozmichuvannya Korpusu ukrajyns'koyi movy. *Ukrajyns'ke movoznavstvo*, 46(1):92–103.
- [7] Darchuk, N. (2013). *Komp'yuterne anotuvannya ukrajyns'koho tekstu: rezul'taty i perspektyvy*. Osvita Ukrainy, Kyiv.
- [8] Derwojedowa, M., Piasecki, M., Szpakowicz, S., Zawislawska, M., and Broda, B. (2008). Words, concepts and relations in the construction of Polish WordNet. In *Proceedings of the Global WordNet Conference*, pages 162–177, Seged, Hungary.
- [9] Erjavec, T. and Fišer D. (2006). Building the Slovene Wordnet: first steps, first problems. In *Proceedings of the Third International WordNet Conference, Vol. 2006*, GWC, Jeju Island, Korea.
- [10] Flanigan, J., Thomson, S., Carbonell, J., Dyer, C., and Smith, N. A. (2014). A discriminative graph-based parser for the abstract meaning representation. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pages 1426–1436, Association for Computational Linguistics, Baltimore, Maryland.
- [11] Haselow, A. (2011). *Typological changes in the lexicon*. 1st ed. Mouton de Gruyter, New York, NY.
- [12] Katsnelson, S. (1972). *Tipologiya yazyka i rechevoe myshlenie*. Nauka, Leningrad.
- [13] Koeva, S., Leseva, S., and Todorova, M. (2006). Bulgarian sense tagged corpus. In *Proceedings of the 5th SALTML Workshop on Minority Languages: Strategies for Developing Machine Translation for Minority Languages*, pages 79–86, Genoa, Italy.
- [14] Kul'chyt's'kyi, I., Romanyuk, A., and Khariv, B. (2010). Rozroblennya Wordnet-podobnoho slovnyka ukrajyns'koyi movy. *Visnyk Natsional'noho universytetu "Lviv's'ka politekhnika"*. *Informatsiyni systemy ta merezhi*, 673:306–318.
- [15] Kustova, G., Lyashevskaya, O., Paducheva Ye., and Rakhilina Ye. (2005). Semanticheskaya razmetka leksiki v Natsionalnom korpusse russkogo yazyka. In *Natsionalnyy korpus russkogo yazyka 2003–2005*, pages 155–174, Moscow.
- [16] Lanhenbakh, M. (2012). *Spoluchuvanist' imennykiv ukrajyns'koyi movy (hlybynnyy semantiko-syntaksychnyy analiz)*. Kyiv.
- [17] Miller, G. A., Beckwith, R., Fellbaum, Ch., Gross, D., and Miller, K. J (1990). Introduction to WordNet: An on-line lexical database. *International journal of lexicography*, 3(4):235–244.
- [18] Palmer, M., Gildea, D., and Kingsbury P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational linguistics*, 31(1):71–106.
- [19] Raffaelli, I., Bekavac, B., Agić, Ž., and Tadić, M. (2008). Building croatian wordnet. In *Proceedings of the Fourth Global WordNet Conference*, pages 349–359, Szeged, Hungary.
- [20] Schenkel, R., Suchanek, F. M., and Kasneci, G. (2007). YAWN: A Semantically Annotated Wikipedia XML Corpus. In Kemper, A., Schöning, H., Rose, T., Jarke, M., Seidl, T., Quix, C., and Brochhaus, C., editors, *BTW. LNI, vol. 103*, pages 277–291, GI.
- [21] Sternin, I. (2015). *Leksicheskoe znachenie slova v rechi*. Direct-Media, Moscow-Berlin.
- [22] Van Valin, R. D. (2001). *An Introduction to Syntax*. Cambridge University Press, Cambridge.
- [23] Tesnière, L. (2015). *Elements of structural syntax*. John Benjamins Publishing Company, Amsterdam – Philadelphia.