

VIOLETTA KOSESKA-TOSZEWA^{1,A}, JOANNA SATOŁA-STAŚKOWIAK¹
& WOJCIECH SOSNOWSKI¹

¹Institute of Slavic Studies, Polish Academy of Science, Warsaw

^Aamaz@inetia.pl

FROM THE PROBLEMS OF DICTIONARIES AND MULTI-LINGUAL CORPORA

Abstract

The article describes the work on a number of dictionaries being developed by the Corpus Linguistics and Semantics Group of the Institute of Slavic PAS. They include “Contemporary Bulgarian-Polish Dictionary”, “Bulgarian-Polish Online Dictionary” and “Russian-Bulgarian-Polish Dictionary”. The dictionaries differ in the numbers of entries, as well as in the different degrees of their connection with parallel corpora being elaborated under the “Clarín” project. All the discussed dictionaries are similar with respect to their use of traditional, syntactic classifiers and of semantic classifiers, introduced for the first time in the existing lexicographical practice. Thanks to the “Polish-Bulgarian-Russian Corpus”, the Group has managed to verify the results of contrasting Polish and Bulgarian in the light of scope-based logical quantification. Thanks to the Russian material added to the trilingual corpus, the researchers have managed to confirm the fact that from the viewpoint of “incomplete quantification” Russian and Polish (synthetic languages) behave similarly, and are opposed to the analytic Bulgarian.

Keywords: Bilingual dictionary, trilingual dictionary, online dictionary, traditional classifier, syntactic classifier, semantic classifier, bilingual parallel corpora, trilingual parallel corpora, linguistic quantification, “incomplete quantification”.

The idea of creating a “Contemporary Bulgarian-Polish Dictionary” is closely related to the collective work of the Bulgarian linguists L. Andrejczin, L. Georgiev, S. Ilczew, N. Kostov, S. Stojkov, S. Todorov (a dictionary of 1973), and especially I. Lekov (1945) and I. Lekov, Fr. Slavski (Eds.) (1961), as well as the greatest Bulgarian-Polish dictionary by Fr. Slavski of 1987. Those outstanding linguistic studies gave rise to the idea of creating a separate work containing about 40,000 entries and introducing the most important, in the authors’ opinion, grammatical and semantic information about Polish and Bulgarian verbs.

The book form of this dictionary will consist of 4 fascicles. The first of them should appear in 2013. As far as possible, the language material will be connected with the “Bulgarian-Polish Parallel Corpus” (authors: L. Dimitrova, V. Koseska-Toszewa) and with the “Polish-Bulgarian-Russian Parallel Corpus” (authors: V. Koseska-Toszewa, J. Satoła-Staškowiak, W. Sosnowski, A. Kisiel) (comments on that subject can be found in publications dating from 2009, 2011 and 2012, see the references). The authors do not take into consideration archaic and dialectal lexis, and try to introduce the maximum possible number of Polish, Bulgarian and Russian neologisms that are already commonly used in the selected languages. Beside traditional classifiers, the authors introduce into the entries new semantic classifiers, which help distinguish between language forms and their meanings. They concern first of all verb forms and adverbial. Such an approach to elaborating those parts of speech is a novum in the existing lexicographic practice.

The task of Joanna Satoła-Staškowiak is to prepare a thesaurus of neologisms, and to introduce the newest lexis into the dictionary. She is responsible for translating the meanings of Bulgarian entries to Polish and, for the correspondence between the Polish and Bulgarian meanings (Dimitrova, Koseska-Toszewa, Satoła-Staškowiak, 2009, 2012; Koseska-Toszewa, Satoła-Staškowiak, Duszkin, 2012).

V. Koseska-Toszewa verifies Bulgarian entries, checks if their translation to Polish is correct, defines semantic classifiers, and is responsible for their introduction into verb entries (Dimitrova, Koseska-Toszewa, 2009; Dimitrova, Koseska-Toszewa, Satoła-Staškowiak, 2009, 2012; Koseska-Toszewa, Satoła-Staškowiak, Duszkin, 2012; Dimitrova, Koseska-Toszewa, 2012), see the article by V. Koseska-Toszewa in this volume, see also the article by J. Satoła-Staškowiak in this volume.

L. Dimitrova is responsible for verifying Bulgarian entries and checking the usage frequency of Bulgarian lexemes based on the “Bulgarian-Polish Corpus” (Dimitrova, Koseska-Toszewa, 2009; Dimitrova, Koseska-Toszewa, Satoła-Staškowiak, 2009, 2012; Dimitrova, Koseska-Toszewa, 2012).

The “Bulgarian-Polish Online Dictionary” is an experimental bilingual dictionary being developed by the authors of the “Contemporary Bulgarian-Polish Dictionary”. The dictionary numbers just about 5,000 entries. The task of preparing a proprietary computer program for handling the electronic Bulgarian-Polish dictionary has been undertaken by L. Panova and R. Dutsova. Besides adding traditional and semantic classifiers, the priority task of the authors of the “Contemporary Bulgarian-Polish Dictionary” in an online version is to include the newest Bulgarian lexis (being developed by J. Satoła-Staškowiak). The most important issue has become the way of examining an individual lexical unit, which consists in establishing whether it is fully accepted by the language system (the determinants in Bulgarian were, for example: the possibility of using the neologism with the article, the fact of possessing a plural form, and others). An appropriate verification of the newest lexemes will protect the authors against the erroneous inclusion in the dictionary of, e.g., occasional expressions, whose life in the system of each language is limited in time.

The plan for developing a Russian-Bulgarian-Polish dictionary was set up two years ago by the Corpus Linguistics and Semantics Group at the Institute of Slavic of the Polish Academy of Sciences. The dictionary is a kind of lexicon containing

short definitions and a large number of meanings in all the three languages. According to the plan, the dictionary will contain about 10,000 entries for each of the three languages, and will number about 30,000 entries altogether. After long discussions, the final order of the described languages was selected. The dictionary starts with Russian due to the largest number of recipients knowing that language; Bulgarian is presented next, and then Polish. The above dictionary concerns languages representing the southern, eastern and western groups of Slavic languages, and will be of a large importance not only for slavists. According to the original plan, its authors were to be Maksim Duszkin, Violetta Koseska-Toszewa, Joanna Satoła-Staśkowiak (on the Polish side) and Anastasja Tzoneva (on the Bulgarian side). For reasons beyond our control, the composition of authors of the trilingual dictionary has been changed. At present, work on the dictionary is carried out by W. Sosnowski, V. Koseska-Toszewa and A. Kisiel. The first issue of the dictionary is to be published in 2014/2015. Examples:

безымя́нн ый <i>adj.</i> 'having no name' <i>безымянная могила</i> <i>безымянный палец</i>	безы́мен ен <i>adj.</i> 'having no name' <i>безыменен гроб</i> <i>безыменен пръст</i>	beziemienn y <i>adj.</i> 'having no name' <i>beziemienny grób</i> <i>palec serdeczny</i>
безысхо́дн ый <i>adj.</i> 'giving no hope for a positive result' <i>безысходное положение</i> <i>see безысходност ь</i>	безы́зход ен <i>adj.</i> 'giving no hope for a positive result' <i>безизходно положение</i> <i>see безызходица</i>	beznadziejn y <i>adj.</i> 'giving no hope for a positive result' <i>beznadziejna sytuacja</i> <i>see beznadziejność</i>
белé ть, -ю, -ешь <i>vi.</i> (<i>state, intransitive</i>) 'become white' <i>небо белеет</i> <i>see побелéть</i>	побеля́в ам <i>vi.</i> (<i>state, intransitive</i>) 'become white' <i>косата на Мария побеля</i> <i>see побелéя</i>	biel eć <i>vi.</i> (<i>state, intransitive</i>) 'become white' <i>niebo bieleje</i> <i>see zbieleć</i>
бели́знá <i>n. f. (Sg. Tantum)</i> 'blinding white colour of sth.' <i>белизна снега</i>	белотá <i>n. f. (Sg. Tantum)</i> 'blinding white colour of sth.' <i>белотата на снега</i>	biel <i>n. f. (Sg. Tantum)</i> 'blinding white colour of sth.' <i>biel śniegu</i>
бел и́ть, -ю, -ишь <i>vi.</i> (<i>state, transitive</i>) 'robić coś białym' <i>белить стены</i> <i>see побели́ть</i>	белóсв ам <i>vi.</i> (<i>state, transitive</i>) 'robić coś białym' <i>белосвам (варосвам) таван</i> <i>избелвам плат</i> <i>see белóсам</i>	biel ić <i>vi.</i> (<i>state, transitive</i>) 'robić coś białym' <i>bielić ściany</i> <i>see pobielić</i>

бѣлк а , <i>gen. plural</i> -лок <i>n. f.</i> 1. ‘small animal from the rodent group with red, black of grey fur and fluffy tail, living on trees’ <i>вертеться или кру-титъся как белка в колесе</i> 2. --- no meaning --- 3. futro lub skóra z tego zwierzęcia i wyrób z niego <i>воротник из белки</i>	кáтериц а <i>n. f.</i> 1. ‘small animal from the rodent group with red, black of grey fur and fluffy tail, living on trees’ <i>въртя се като дявол на шми; не подвигам крак, трепя се от сутрин до вечер</i> 2. --- no meaning --- 3. --- no meaning ---	wiewiórk а <i>n. f.</i> 1. ‘small animal in the rodent group with red, black of grey fur and fluffy tail, living on trees’ <i>kręcić się jak w kołowrotku</i> 2. <i>coll.</i> ‘about a person having red hair’ 3. --- no meaning ---
белкóв ый <i>adj.</i> ‘built of proteins’ <i>белковое вещество</i>	белтѣч ен (сѣстав) <i>adj.</i> ‘built of proteins’ <i>белтъчен състав</i>	białkow y <i>adj.</i> ‘built of proteins’ <i>substancja białkowa</i>

As we have already mentioned, our dictionaries will employ syntactic classifiers with the abbreviations: *transitive* and *intransitive*, which indicate transitivity and intransitivity of a verbum. By **transitivity** we will mean the possibility of a direct object appearing in the sentence after the verb. In Polish and Russian sentences, a transitive verb is followed by nomen in the accusative case. **Intransitivity** excludes the above possibility. Intransitive verbs cannot be followed by a direct object, and in Polish and Russian the verbum form is followed by all case forms except the accusative case. As our dictionary is not a valence dictionary, the above definition should be satisfactory for the reader.

Yet another group of classifiers will be **semantic classifiers**, which were not elaborated earlier, and were distinguished thanks to the many years of work on the first semantic confrontation in the word, carried out in the 12-volume monograph *Contrastive Polish and Bulgarian Grammar* with a semantic interlanguage (Koseska-Toszewa, 2006). The authors are aware that distinguishing semantic classifiers is not an easy task, and in order to achieve such a goal, one should **consistently distinguish between the form and the meaning**, e.g. decide whether the imperfect or perfect aspect used as a classifier in the commonly known dictionaries is a form of the verb or its meaning? In the authors’ opinion, this a form, and its meanings are either 1) states or sequences of states and events ending with a state, or 2) events or sequences of states and events ending with an event.

The notions of events, states and their configurations are understood here as in the **network-based description of time and aspect**, i.e. so that an event does not last in time (it begins, ends or interrupts states, while a state lasts and is begun or ended by an event, see in more detail the use of Petri net theory and its application in a natural language (Mazurkiewicz, 1986; Koseska-Toszewa, A. Mazurkiewicz, 1988, 2010; Koseska-Toszewa, 2006). Assuming that a verb form marked as “imperfective” is a form, we will present its meanings as either 1. *state*

— a state or 2. — a sequence of events and states ended with a state, while a verb form marked as “perfective” will have a semantic classifier either 1. *event*, i.e. an event, or 2. a sequence of states and events ended with an event. Meanings 1 and 2 for an *event* and 1 and 2 for a *state* can be clearly illustrated on the example of an aspectually-temporal relation, i.e. when the verb form expresses a specific tense, due to which the infinitive form, which is “timeless”, can only be accompanied by the *state* and *event* abbreviations. Examples:

(Rus.) **писа́ть** *vi.* (*state, transitive*) (Bulg.) **пи́ша** *vi.* (*state, transitive*) (Pl.) **писа́ć** *vi.* (*state, transitive*)
 (Rus.) **написа́ть** *vp.* (*event, transitive*) (Bulg.) **напи́ша** *vp.* (*event, transitive*)
 (Pl.) **писа́ć** *vp.* (*event, transitive*)
 (Rus.) **жале́ть** *vi.* (*state, transitive*) (Bulg.) **съжаля́вам** *vi.* (*state, transitive*)
 (Pl.) **żałować** *vi.* (*state, transitive*)

Semantic classifiers concerning quantification of nomen and predicates are initially introduced in the “Polish-Bulgarian-Russian Parallel Corpus” (the corpus, being developed by the authors, is created within the European Project Clarin¹). The authors transfer language material from that parallel corpus to the planned trilingual dictionary.

The discussed three dictionaries differ in the numbers of entries. The *Bulgarian-Polish Online Dictionary* contains a much smaller number of entries than the *Contemporary Bulgarian-Polish Dictionary* or the *Russian-Bulgarian-Polish Dictionary*, which follows from the fact that it is treated as a scientific experiment by the authors. However, there is a possibility of extending the language resources of the dictionary. The authors are planning to make use of this option, and introduce supplements in the subsequent years. The second largest dictionary is the *Russian-Bulgarian-Polish Dictionary*, which includes about 30,000 entries. The largest dictionary in the cycle is the *Contemporary Bulgarian-Polish Dictionary*, having about 40,000 entries.

The dictionaries differ also in the number of semantic classifiers. The *Contemporary Bulgarian-Polish Dictionary* is equipped in semantic classifiers concerning solely verbum. This is also the case with the *Bulgarian-Polish Online Dictionary*, where only verbum is characterized with help of semantic classifiers. We are planning that in the *Russian-Bulgarian-Polish-Dictionary* semantic classifiers will only concern verbum, but a classifier with quantification of adverbia will also be introduced in all three languages.

All the three dictionaries described in the article are connected with parallel corpora, and can be called “contemporary” due to the emphasis placed on the newest

¹(Common Language Resources and Technology Infrastructure) — a scientific project, which in February 2012 obtained the European Commission legal status of ERIC (European Research Infrastructure Consortium). The founders of Clarin ERIC are Austria, Bulgaria, the Czech Republic, Denmark, Estonia, Germany, the Netherlands and Poland. CLARIN is a project from the so-called ESFRI road map (European Roadmap for Research Infrastructures, European Strategy Forum on Research Infrastructures). The main goal of the project is to combine *language resources and tools* for European languages into a single, common and uniform network, which is to become an important working tool for scientists from the humanities in the broad sense of the term.

and contemporary lexis in the language material under preparation, as well as due to the analysis of the “usage frequency” verified by the parallel corpora mentioned here. In lexicographic work, the most problems are posed by the trilingual Russian-Polish-Bulgarian dictionary due to the methodological assumption concerning the entries elaborated in all languages, saying that the material should be described in a way enabling unproblematic establishment of a database for future dictionaries, also electronic ones. This means, among others, that the entries should contain meanings and classifiers in each of the languages rather than as is the case in traditional bi- and trilingual dictionaries, where the meanings and classifiers concern the initial language only.

By way of example, the semantic annotation applied results from a very complicated process, see the article by V. Koseska-Toszewa in this volume. For this reason, we could limit it to marking just the quantification of nomen **at the subject position**, and to specifying **events** — **event 1**, **event 2** (sequence of events and states, finally ended with an event) and **states** — **stat1** and **stat2** (sequence of states and events, finally ended with a state). It is also worthwhile to distinguish between adverbial forms and their quantifying meanings, see the table with „zawsze” — adverbium with the universal quantificational meaning — given below.

By **existentiality** we mean here expressions of the form $(\exists x)P(x)$ preceding the predicate, i.e. a sentential function P (from now on, P), in the semantically-logical structure of the sentence, read using the phrases “there is an x , such that”, “for some x ”. By **universality** we mean expressions of the form $(\forall x)P(x)$ preceding the predicate P in the semantically-logical structure of the sentence. Finally, by **uniqueness** we mean an expression of the form $(ix)P(x)$ which assumes that the given sentential function (P) is satisfied either by exactly one element of the considered universe, or by one and only one set of elements. As generally accepted in logical literature by now, we treat the iota operator as the unique quantifier. Quantificational expressions are not unambiguous (Koseska, Gargov, 1990). For example, “each such that P ” may be understood as “all elements satisfying P ”, which we can write in the form $(\forall x)P(x)$. See *Днес всяко момче кара ски.* / *Dzisiaj każdy chłopiec jeździ na nartach.* [Today each boy is skiing]. In a context different from the above, the expression written as $(\forall x)P(x)$ can be understood as “the set (of generally many elements) satisfying P as the only one”. Then this quantificational meaning should be written as $(iX)P(X)$, and it would represent an expression with a unique meaning. When encountering “incomplete quantification”, we precede the quantification symbols with the question mark “?” — e.g. $?(ix)Px$ — incomplete unique (iota-operator) quantification, see below the quantificational meanings of adverbia referring to time and aspect of verbs. This is why we have to do with quantification of the predicate rather than nomen here.

Всегда	Винаги	Zawsze
adv. $(\forall X) P(X)$, univ.	adv. $(\forall X) P(X)$, univ.	adv., $(\forall X)P(X)$, univ.

See also (Tab. 1) semantic annotation taking into consideration quantification of names at the subject position, and the aspectually-temporal meaning of the Polish, Bulgarian and Russian verbs in the Bulgarian fable translated to Polish and Russian *Цар Безсънко* from the Polish-Bulgarian-Russian Parallel Corpus.

Table 1.

Król Bezsenek	Цар Безсънко	Царь Бессон
<i>Był sobie</i> $(\exists X)P(X)$ (state) krasnoludek $?(\exists x)P(x)$ z długą białą brodą w czerwonej czarpcie.	<i>Имаше някога</i> $(\exists X)P(X)$ (state) едно джудже $(\exists x)P(x)$ с дълга бяла брадица и алена шапчи- ца.	<i>Жил некогда</i> $(\exists X)P(X)$ (state) на свете гном $?(\exists x)P(x)$ с длинной бе- лой бородой.
$(\exists X)P(X)$ (event) <i>Pew- nego wiosennego poran- ka siadł</i> $?(ix)P(x)$ kras- noludek jak zwykle przed swoim domkiem i począł liczyć ziarenka w swoim różańcu.	$(\exists X)P(X)$ <i>Една про- летна утрин</i> $(ix)P(x)$ (event) нашето джуд- же излезе както винаги пред вратата на своята къщица и започна да върти броеничката си от просени зрънца.	$(\exists X)P(X)$ (state) <i>Как- то весенним утрлом</i> $?(ix)P(x)$ гном по обыкновеню <i>сидел</i> пе- ред своей избушкой и перебирал четки.
$?(ix)P(x)$ Myszka $(iX)P(X)$ <i>przyglądała</i> (state) mu <i>się</i> bacznie błyszczącymi oczyma, oddychając ciężko.	$(ix)P(x)$ Мишката го <i>гледаше</i> (state) с малки- те си лъскави очички, без да мига, и дишаше тежко.	$?(ix)P(x)$ Мышка $(iX)P(X)$ (state) <i>смотре- ла</i> на него маленькими блестящими глазами и тяжело дышала.
$?(ix)P(x)$ Krasnoludek $(iX)P(X)$ (event) <i>przerwał</i> <i>śpiew</i> i wstał, aby przy- witać nieznanego goś- cia.	$(ix)P(x)$ Джуджето $(iX)P(X)$ (event) <i>преста- на да пее</i> и се изправи на нозете си.	$?(ix)P(x)$ Гном $(iX)P(X)$ (event) <i>перестал петь</i> и встал.

Attention! The symbol "?" preceding logical symbols related to quantification represents incomplete quantification.

The material from our three-lingual parallel corpus reveals the fact, known from the „Bulgarian-Polish Contrastive Grammar”, that synthetic languages differ from analytic ones in having language means expressing especially unique quantification omitted in them, see (Koseska, Gargov, 1990; Koseska, 2006; Koseska, Korytkowska & Roszko, 2007). Let us remind here that, following Ajdukiewicz, such a phenomenon is referred to as “incomplete quantification” (Ajdukiewicz, 1974). As he wrote in his “*Logika pragmatyczna*”: “...an expression is called, for its certain meaning, a statement (or a sentence in the logical sense) if the expression, with just this meaning, utters some judgment, i.e. some thought referring in a reporting way to a certain state of things”. Ajdukiewicz called statements sentences in the logical sense. In his opinion, they had a certain common property distinguishing them from all other types of sentences. “Namely, all statements (and only them!)

are either true or false.” However, in a natural language we often encounter statements without either the true or the false value, such as, e.g., the sentence from “*Logika pragmatyczna*” by Ajdukiewicz: *Anglicy są flegmatyczni* [The English are phlegmatic]. Ajdukiewicz draws the reader’s attention to the fact that the above sentence represents incomplete quantification, which is why it has neither the true or the false value. Indeed, we do not know whether it refers to all the English, to their majority, or only to some of the English. Roughly speaking, the sender of the above information does not know whether the predicate “x are phlegmatic” is satisfied for the name “the English”. In this case, we need additional information, which can give the true value to this predicate. Namely, the information about scope-based quantification of the name “the English” is missing here. Ajdukiewicz assessed this phenomenon as error in articulation. Omitting of some component in a sentence prevents judging it as either true or false, and leads to misunderstanding. Incomplete quantification or relativization was deemed to lead to misunderstanding the speaker’s intention or to corrupting it. Clearly, a natural language has an effective defence against such a situation. Whoever does not understand the information sender’s intention, asks them a supplementary questions, and in the end clarifies any misunderstanding. However, there is still the question if “incomplete quantification”, which is the subject of our observations, is indeed a marginal problem — an ordinary error? Observations of research on the semantic category of definiteness / indefiniteness in Polish and Bulgarian show that incomplete quantification cannot be an ordinary language mistake, since languages differ among themselves in this respect. In Polish we encounter such as phenomenon more often than in Bulgarian. Ajdukiewicz spoke of incomplete quantification having in mind the nominal phrase only, while we speak of incomplete logical scope-based quantification having in mind also the verbal phase in sentences of both the contrasted languages.

The languages studied earlier — Polish and Bulgarian — differ from each other with respect to incomplete quantification. This fact can be explained by differences in the morphological systems of both languages. In Polish this is a common phenomenon. Incomplete unique quantification is limited in both languages, but in Bulgarian it is not encountered on the nominal phrase level first of all because Bulgarian is an article language, see (Koseska, Gargov, 1990, p. 138–139), e.g.:

Pl. *Odwiedziłem staruszkę* (This only one, some, certain?) — incomplete quantification of the nominal phrase.

Bulg. *Посетих старуцама*, but it is impossible to build the sentence *Посетих старица with the same meaning. / *Odwiedziłem tę jedyną staruszkę* [I visited this unique old woman] (unique quantification).

Посетих една старица / *Odwiedziłem pewną staruszkę* (existential quantification of the nominal phrase).

The examples from our corpus show that Russian behaves from the viewpoint of incomplete quantification just like Polish, and that synthetic languages (Polish and Russian) are in opposition to the analytic Bulgarian language with respect to „incomplete quantification”. See below examples isolated in this respect from

the “Polish-Bulgarian Russian Parallel Corpus”, which is being developed by the authors of this article under the European „Clarin” grant.

?(ix)P(x) Myszka (iX)P(X) (state) <i>przyglądała mu się bacznie błyszczącymi oczyma, oddychając ciężko.</i>	(ix)P(x) Мишката го (iX)P(X) (state) <i>гледаше с малките си лъскави очички, без да мига, и дишаше тежко.</i>	?(ix)P(x) Мышка (iX)P(X) (state) <i>смотрела на него маленькими блестящими глазами и тяжело дышала.</i>
?(ix)P(x) Krasnoludek (iX)P(X) (event) <i>przerwał śpiew i wstał, aby przywitać nieznanego gościa.</i>	(ix)P(x) Джуджето (iX)P(X) (event) <i>престана да пее и се изправи на нозете си.</i>	?(ix)P(x) Гном (iX)P(X) (event) <i>перестал петь и встал.</i>

Authors of the article are aware that the plans discussed here require many years of work of the small Corpus Linguistics and Semantics Group of IS PAN.

References

- Ajdukiewicz (1974). *Logika pragmatyczna*, Warsaw.
- Андрейчин, Л., Георгиев, Л., Илчев, С., Костов, Н., Леков, И., Стойков, С., Тодоров, Ц., (1973). *Български тълковен речник*. Наука и изкуство. III издание.
- Bogusławski, A. (2003). *Aspekt i negacja*, Warsaw.
- Dimitrova, L., Koseska-Toszeva, V. (2009a). Bulgarian-Polish Corpus. *Cognitive Studies / Études Cognitives*, 9, p. 133–141.
- Dimitrova, L., Koseska-Toszeva, V. (2009b). Classifiers and Digital Dictionaries. *Cognitive Studies/Études Cognitives*, Vol. 9, p. 117–131.
- Dimitrova, L., Koseska-Toszeva, V., Dutsova, R., Panova, R. (2009). Bulgarian-Polish online Dictionary — Design and Development. In: Koseska, Dimitrova & Roszko (Eds.), *Representing Semantics in Digital Lexicography. Proceedings of the MONDILEX Fourth Open Workshop, 29 June – 1 July, 2009, Warsaw*, SOW Publishers, p. 76–88.
- Dimitrova, L., Koseska-Toszeva, V., Satoła-Staškowiak, J. (2009). Towards a Unification of the Classifiers in Dictionary Entry. In: Garabík (Ed.), *Metalanguage and Encoding Scheme Design for Digital Lexicography. Proceedings of the MONDILEX Third Open Workshop, 15–16 April, 2009, Bratislava*, Slovakia, p. 48–58.
- Dimitrova, L., Panova, R., Dutsova, R. (2009). Lexical Database of the Experimental Bulgarian-Polish Online Dictionary. Garabík (Ed.), *Metalanguage and Encoding scheme Design for Digital Lexicography. Proceedings of the MONDILEX Third Open Workshop, 15–16 April, 2009, Bratislava*, p. 36–47.
- Dimitrova, L., Koseska-Toszeva, V. (2012). Bulgarian-Polish Parallel Digital Corpus and Quantification of Time. *Cognitive Studies/Études Cognitives*, 12, p. 199–208.
- Dimitrova, L., Koseska-Toszeva, V., Satoła-Staškowiak, J. (2012). Neologisms in Bilingual Digital Dictionaries (on the example of Bulgarian-Polish Dictionary). *Cognitive Studies/Études Cognitives*, 12, p. 199–208.
- Dimitrova, L., Dutsova, R. (2012). Implementation of Bulgarian-Polish Online Dictionary. *Cognitive Studies/Études Cognitives*, 12, p. 219–229.

- Feleszko, K., Koseska-Toszewa, V., Sawicka, I. (1974). Związki aspektu z temporalnością w językach południowosłowiańskich. In: *Studia z Filologii Polskiej i Słowiańskiej*, XIV, p. 183–187.
- Garabík, R., Dimitrova, L., Koseska-Toszewa, V. (2011). Web-presentation of bilingual corpora (Slovak-Bulgarian and Bulgarian-Polish). *Cognitive Studies/Études Cognitives*, 11, p. 227–239.
- Gołąb, Z., Heinz, A., Polański, K. (1968). *Słownik terminologii językoznawczej*, Kraków.
- Karolak, S. (2008). O kompozycjonalnym ujęciu aspektualnej struktury verbum. In: *Językoznawstwo historyczne i typologiczne: w 100-lecie urodzin profesora Tadeusza Milewskiego*, series: Rozprawy Wydziału Filologicznego / Polska Akademia Umiejętności; vol. IV, Kraków, p. 279–287.
- Koseska-Toszewa, V. (1974). Z problematyki temporalno-aspektowej w języku bułgarskim (Relacja imperfectum — aoryst), *SFPS*, XIV, p. 213–226.
- Koseska-Toszewa, V. (2006). Semantyczna kategoria czasu. In: *Gramatyka konfrontatywna bułgarsko-polska*. Sławistyczny Ośrodek Wydawniczy Publishers, Warsaw.
- Koseska-Toszewa, V., Gargov, G. (1990), *Bułgarsko-polska sâpostavitelna gramatika*, vol. 2. *Semantichnata kategorija opredelenost/neopredelenost*, Sofia.
- Koseska-Toszewa, V., Mazurkiewicz, A. (1988). Net Representation of Sentences in Natural Languages. In: *Lecture Notes in Computer Science 340. Advances in Petri Nets 1988*, Springer-Verlag, p. 249–266.
- Koseska-Toszewa, V., Korytkowska, M. & Roszko, R. (2007). *Polsko-bułgarska gramatyka konfrontatywna*, Wydawnictwo Akademickie Dialog Publishers, Warsaw.
- Koseska-Toszewa, V., Mazurkiewicz, A. (2010). *Time flow and Tenses*, SOW Publishers. Warszawa.
- Koseska-Toszewa, V., Satoła-Staškowiak, J., Duszkin, M. (2012). Теоретические и методологические аспекты создания польско-булгарско-русского словаря, In: *Mega-Ling'2011. Горизонти прикладної лінгвістики і лінгвістичних технологій*, Національна Академія Наук України Український Мовно-Інформаційний Фонд, p. 210–219.
- Koseska-Toszewa, V., Satoła-Staškowiak, J., Duszkin, M. (2012). Polish-Bulgarian-Russian, Bulgarian-Polish-Russian or Russian-Bulgarian-Polish Dictionary? *Cognitive Studies/Études Cognitives*, 12, p. 227–239.
- Koseska-Toszewa, V., Satoła-Staškowiak, J. (2013). O pracy nad słownikami książkowymi i elektronicznymi z językiem bułgarskim. In: *Сборник с материали от Шестата национална конференция с международно участие по лексикография и лексикология. (70 години булгарска академична лексикография, София, 24–25 октомври 2012)*. София, АИ Проф. М Дринов. (in print)
- Lekov, I. (1945). *Полско-булгарски речник*, BAS Publishing House, Sofia. Bulgaria.
- Lekov, I., Sławski Fr. (Eds.) (1961). *Полско-булгарски речник*, BAS, Publishing House, Sofia. Bulgaria.
- Mazurkiewicz, A. (1986). Zdarzenia i stany: elementy temporalności, In: V. Koseska-Toszewa, I. Sawicka, J. Mindak, (Eds.), *Studia gramatyczne bułgarsko-polskie*, Vol. 1: *Temporalność*, Wrocław, p. 7–21
- Satoła-Staškowiak, J. (2010). *Polsko-bułgarskie odpowiedniości przekładowe czasów przeszłych*, Sławistyczny Ośrodek Wydawniczy, Warsaw.
- Sławski, F. (1987). *Podręczny słownik Bułgarsko-Polski z suplementem*, Państwowe Wydawnictwo Wiedza Powszechna. Warszawa.