



SHARED BY ALL SPEAKERS? DATIVE PREDICATIVES IN BULGARIAN AND RUSSIAN

Elena Ivanova

St. Petersburg State University (Russia)

Anton Zimmerling

*Pushkin State Russian Language Institute / Moscow Pedagogical State University / Institute of
Linguistics, Russian Academy of Sciences (Russia)*

Abstract. We analyze the language-internal variation in the class of dative predicatives in Bulgarian and Russian. For each language, a test questionnaire is prepared. The stimuli are grouped into 15 thematic classes for Russian and 19 thematic classes for Bulgarian and tested on native speakers. Their responses provide a sample of stimuli ranked according to the approval rate. The same set of stimuli is tested on Russian National Corpus and Bulgarian National Corpus, which provided the second ranked sample. The volume and structure of the class of dative predicatives is established by the ratio of these ranked samples.

Keywords: Russian; Bulgarian; corpus linguistics; sociolinguistics; variation; lexicon; grammar; predicatives

Slavic languages have a class of non-verbal predicatives expressing state and modality. A subclass of Slavic predicatives selects dative subjects, cf. Russ. *Мне скучно / зябко / сонно / простительно / невдомек*, Bulg. *Любопитно / празнично / гузно / спешно / излишно / позволено ми е*. Despite the class of elements licensing dative-predicative structures (hence — DPS) is open in many Slavic languages including Russian, Bulgarian, Ukrainian, Serbian, Slovenian lexicographic descriptions of DPS predicatives based on representative corpus data until recent time have been lacking in Bulgarian and Russian studies.

The undertaken contrastive analysis of Bulgarian and Russian DPS predicatives aims at describing the nucleus of the DPS construction and the mechanisms of its extension. The procedure follows the double sociolinguistic vs corpus analysis of Russian DPS (Zimmerling, 2017; 2018a). A model of the DPS construction has been build. This model predicts that the DPS construction has a nucleus {Lex_k} which belongs to the lexicon and an extension which belongs to grammar {Gram_k}. The main hypothesis is that the speakers apply to the same principles of semantic

selection and add new elements to $\{\text{Gram}_k\}$ even if they use non-identical inventories of lexical items. A double sociolinguistic/ corpus study of Bulgarian DPS based on this procedure was carried out in 2017-2018 by Elena Ivanova (Ivanova, 2018). The distribution of DPS predicatives in other Slavic languages has been initiated in (Maric & Kerkez, 2018; Mitkovska 2018; Uhlik 2018; Kulinich 2018; Petrova 2018), the results have been presented at the thematic panel on DPS predicatives at the XVI International Slavic Congress in Belgrade (2018)¹.

Below follows the description of the undertaken sociolinguistic and corpus analysis of Russian and Bulgarian data and a discussion of its results in a contrastive perspective.

1. Research program

The research is based on the following procedure.

1. We assume that the DPS construction in Bulgarian and Russian has one and the same basic semantics — the meaning of inner state linked with a referential animate subject during a period of time (Zimmerling, 2018b).

2. The DPS construction is used in a number of denotative situations. These situations constitute the DPS ontology. The invariant meaning of inner state combines with different types of the denotative situation.

3. DPS predicatives in both sets of stimuli are grouped in 15 thematic classes which represent the standard variant of the DPS ontology. Additional thematic classes are added if needed.

4. The frequency rates for the same sets of stimuli are tested on corpus data.

5. The volume and structure of the class of DPS predicatives is established by the ratio of two ranked samples. The first sample is ranked according the approval rate of the test stimuli by the native speakers. The second ranked sample pertains to the frequency rates of the same set of stimuli in the corpus.

2. The derivation of DPS predicatives in Russian and Bulgarian

Bulgarian and Russian DPS sentences have a similar build. The Bulgarian DPS construction has three obligatory elements which must be realized overtly: the predicative, the auxiliary in 3Sg and the pronominal clitic in the dative case (with a minor group of predicatives — in the accusative case). In Russian, the BE-auxiliary is dropped in the present indicative. Cf. Bulg. *Неприятно му беше; Мъчно ли ти е?; Страх ме е;* Russ. *Ему было неприятно; Тебе грустно?; Мне страшно.*

The Russian DPS construction is contained by the contrast of three types of word-building stems. Type I stems dubbed ‘stems of actant polarity’ in (Zimmerling, 2010) produce the names of properties (*злой X, гневный X*) and do not produce the names of situations, hence the ill-formedness of stage-level predicates (**Мне злобно, *Мне гневно*). Type II stems dubbed ‘stems of situational polarity’ produce the names of situations expressed by the DPS predicatives (*X-у стыдно, совест-*

но) and do not produce the name of properties, hence the absence of adjectives *стыдный, *совестный in Modern Russian. Type III stems are ambivalent and produce both names of properties and names of situations, hence the parallel uses of the agreeing adjectival and non-agreeing predicative forms with the –o-final: Он груст-ен – Ему грустн-о (весел-ый, скверн-ый; весел-о, скверн-о etc.).

Bulgarian lacks Type II stems with the –o –final. Type II stems with situational polarity is only represented by few predicatives of non-adjectival origin, cf. жал, еня. Type I stems i.e. stems with actant polarity (cf. гладен, пиян, влюбен) does not produce DPS predicatives in standard Bulgarian, but the web attests occasional uses like ?гладно ми е, ?пияно ми е, ?влюбено ми е. The vast majority of the DPS uses represent Type III ambivalent stems, hence the parallel uses of agreeing adjectives and non-agreeing DPS elements: топъл – топло ми е, тежък – тежко ми е, задушен – задушно ми е (Gradinarova, 2010: 34 – 35; Gradinarova, 2017: 64 – 65).

The Bulgarian DPS construction is fed by a broader variety of derivational sources than the Russian DPS construction (Gradinarova, 2018). It licenses several elements with participial morphology and facilitates the transition from Type I stems to Type III stems by licensing derived DPS uses like гордо ми е, виновно ми е, приказливо ми е. Bulgarian has a higher percentage of DPS elements with a nominal morphology (мъка ми е, мерак ми е) and, unlike Russian, licenses a limited number of predicatives with an accusative marking on the animate subject (яд ме е, срам ме е, страх ме е). Meanwhile, Russian has more DPS elements with adverbial and pronominal morphology including fossilized formations like X-у недосуг, нельзя, впору, влом, негоже, нипочем, некстати, поделом.

Both Russian and Bulgarian have DPS elements with the inner structure of a prepositional phrase, but the overall number of such elements is higher in Russian. Cf. Bulg. по път ми е; по джоба, не ми е до..., Russ. X-у не по себе, без толку, не к спеху, в диковинку, не по нраву, к лицу, по фигуре, по размеру, в самый раз etc.

3. Thematic classes of DPS predicatives and the questioning of Russian and Bulgarian speakers

The test questionnaire for Russian includes 422 stimuli. The test questionnaire for Bulgarian includes 320 stimuli. The lists of DPS items are not exhaustive. The undertaken study was aimed at explaining how native speakers add new elements to the DPS class.

The results of the sociolinguistic experiment undertaken in (Zimmerling, 2017) suggest that Russian speakers reproduce the DPS construction by a mixed strategy: they borrow the shared DPS nucleus from the lexicon and apply the rules of semantic selection (s-selection) to an open class of predicative elements including occasional and potential formations. The rules of s-selection in the specified sense can be identified with grammar, so that the reproduction of DPS crucially depends

on the interaction of lexicon and grammar: one part of the DPS class in the active vocabulary of an average speaker comes from the lexicon, while the other part is produced on the basis of the principles of Russian grammar presumably shared by all or most speakers. Similar hypotheses can be tested on other Slavic languages including Bulgarian.

The questionnaire for Bulgarian DPS items was subdivided into the main and additional lists. The main list includes 247 stimuli grouped into 15 thematic classes based on the same DPS ontology as in the Russian experiment (Zimmerling, 2018a):

- 1) Physical state (28 items): *Задушно / хладничко / сънливо ми е; жегга ми е*
- 2) Modality (16): *Невъзможно / простено / немислимо ми е; време ми е*
- 3) Emotional state (67): *Обидно / неспокойно / тегаво ми е; жал ми е*
- 4) Moral evaluation (9): *Гузно / неловко ми е; хак ми е*
- 5) Comfort (6): *Удобно / комфортно ми е; по път ми е*
- 6) Aptitude/inaptitude (2): *Неуместно / нелепо ми е*
- 7) Internal need (2): *Спешно / потребно ми е*
- 8) Compliance to the task (9): *Привично / присъщо / скъпо ми е; по джоба ми е*
- 9) Performance ease (7): *Тежко / лесно / непосилно ми е*
- (10) Willingness to perform (5): *По сърце ми е; не ми е до X*
- (11) General assessment (27): *Нормално / непоносимо ми е; супер / кеф / гот / ок ми е*
- (12) (Ir) relevance (6): *Важно / безразлично ми е; все едно ми е*
- (13) Efficiency (4): *Изгодно / вредно / здравословно ми е*
- (14) Sensoric and intellectual reaction (20): *Съмнително / глупаво / безинтересно / странно ми е*
- (15) Parametric property (39): *Шумно / късно / просторно / рано ми е*

The additional list includes 4 classes of Bulgarian DPS predication added on the basis of semantic and morphosyntactic criteria.

(16) States of mind indicating a symptomatic activity of the animate subject (8): *ядно* ‘X is in the evil mind’, *заядливо* ‘X quibbles’, *сръдливо* ‘X behaves peevishly’, *бъбриво* ‘X got talking’: *Днес ми е много **бъбриво*** ‘I am in a mood of chatting a lot today’ (Gradinarova, 2017: 82 – 83).

(17) Lexicalized participles (9) like *отблъскващо* ‘repulsive’, *потискащо* ‘overwhelming’, ‘oppressing’, *приповдигнато* ‘on the rise’, lit. ‘lifted’: *Тепърва се сблъскам с този вид продажби и малко ми е **объркващо*** ‘This is the first time I’ve come across this type of sale and am a bit **confused**’ (lit. ‘to-me is a bit confused’), *Изнервено ми е и явно имам нужда от въздух* ‘I am all nerves. Apparently, I need fresh air’; *Смачкано ми е след нашия разговор* ‘I feel **depressed** after our conversation’ (lit. ‘to-me is depressed’).

(18) Quasi-emotives i.e. metaphoric transpositions of the names of properties and outer states into the sphere of the animate subject (28). Bulgarian social media frequently apply to the transposed uses of color metaphor (*сиво, синьо ми е*), state

of nature (*дъждовно, облачно ми е*) or produce occasional DPS derivatives from the names of human properties (*сентиментално, оптимистично, флегматично ми е*) (Gradinarova, 2018; Petrova, 2018).

19) Predicatives with accusative marking on the subject (12). Different authors provide non-identical lists of such elements. We decided on checking the following items: *страх* 'fear', *срам* 'shame', *яд* 'anger', *гнус* 'disgust', *грижа* 'care', *еня* 'care', *мързел* 'laziness', *гъдел* 'tickle', *грях* 'guilt', *гняв* 'rage, anger', *смях* 'laughter', *студ* 'cold, chill'.

4. The sociolinguistic experiment: the description and results

We questioned 18 native speakers of Russian, women and men from 18 до 65 years, and 19 native speakers of Bulgarian, women and men from 16 to 65 years. The speakers evaluated the stimuli from the test sets in the frame of the structure without subject-predicate agreement and with an overt subject element in the oblique case. The speakers were not informed about the goal of the experiment.

Ranking of the stimuli. If a DPS predicative got 19 positive responses from all 19 Bulgarian speakers, we assigned it the highest rank '1'. If all speakers rejected the stimulus (0 positive responses), we assigned it the lowest rank '20'. The approval value on the scale 0 20 was dubbed '*Socio* rate', using the terminology of (Zimmerling, 2017).

34 stimuli (10,6% of the whole test set) have the highest rank '1' which confirms their status as standard lexical items in the active vocabulary of all speakers. The Russian experiment provided similar figures, whereby the highest rank was assigned to 14% of the main list stimuli (342 items).

126 Bulgarian stimuli (ca. 40 %) are located in the range 1 £ Socio £ 6. This group includes the DPS elements approved by more than two thirds of the speakers. We tentatively interpret it as {Lex_k} i.e. as the lexical nucleus of the Bulgarian DPS construction, with the lower limit Socio = 6.

44 Bulgarian stimuli (13,8%) are located in the mid-range 9 £ Socio £ 12. The corresponding figure for Russian was 18, 7% (64 stimuli). This group of DPS predicatives can be identified as {Gram_k} i.e. the grammatical extension of the DPS class. The tentative lower limit Socio = 12 is a mere estimate, but it seems natural to conclude that a stimulus does not belong to the shared lexicon if it gets one third of negative responses from the speakers (Zimmerling, 2017: 474).

The lower part of the range (13 £ Socio £ 19) comprises occasional formations (*симетрично, смело, гордо, инфантилно*) and archaisms including several predicatives with accusative case marking (*смях, студ*).

From 320 Bulgarian stimuli, more than one half (164 items, 51%) was approved the half of the speakers or more (Socio ³ 10). The corresponding figure for the main list of the Russian stimuli is considerably higher (71,6%). This contrast is partly due

to the fact that most substandard Russian DPS predicatives, fillers and occasional formations with the low expectancy degree were excluded from the main list (322 items) and included in the additional list (80 items), while this has not been done in the Bulgarian set of stimuli.

In order to get a control measure, we calculated the median value for the idiolects. For Russian, it is set by I_{10} , i.e. by the use of the Russian speaker who has 71,3% of the DPS stimuli in his active vocabulary. The corresponding median value for Bulgarian is lower — the Bulgarian idiolect I_3 has 56, 3% of the whole set of the DPS stimuli, but the retrieved data is in accord with the overall percentage of the stimuli approved by more than one half of the speakers: 71,6% for the main list of the Russian stimuli and 51% for the main list of the Bulgarian stimuli. Thus, both measures — the median value and the *Socio* — give similar results for both languages.

The idiolects of the speakers. The Tab.1 gives the summary of the use of the DPS items by the Russian speakers $I_1 - I_{18}$. The Tab.2 gives the summary on the Bulgarian speakers. The age of the speakers (the second line from the top) is given by the time of the experiment (March 2016 for the Russian speakers and February-April 2018 for the Bulgarian speakers).

The idiolects with the extreme lowest and highest figures were excluded from further processing. The volume of the DPS class by the remaining 17 Bulgarian speakers is located in the range from 33,4% (I_1) up to 68,8% (I_5), which corresponds to 107 up 220 DPS items. The variation in the active DPS vocabulary by the Russian speakers is equally salient: from 169 up to 306 main list DPS items.

Interpretation. The experiment proved that the class of DPS predicatives is not taken from lexicon in its entirety but is partly rebuild by every speaker on the basis or grammatical rules.

The ranking of the stimuli shows that all speakers distinguish the stable nucleus $\{Lex_k\}$ inherited from the lexicon and the varying part. 193 stimuli (over 60%) in the Bulgarian experiment got one third and more negative responses in the range 7 £ Socio £ 19. These elements are identified as $\{Gram_k\}$ i.e. grammatical extension of the DPS construction. The compatibility of marginal DPS elements must be checked for every idiolect on its own basis.

The lower median of the Russian sample $I^m = 71,3\%$ corresponds to the level of 244 DPS stimuli of the main set. The *Socio* gives an almost identical figure — 245 main list items: this number corresponds to the upper part of the ranked sample ($Socio^3 9$) approved by more than 50% of the speakers. For Bulgarian, the median value corresponding to the level of 180 DPS items (for the whole set of stimuli) is exemplified by the idiolect I_3 in Tab.2.

5. The corpus analysis of Bulgarian DPS predicatives

The second part of our research involved the testing of the same set of the stimuli on Russian and Bulgarian corpora — RNC (23 803 881 sentences, checked

Tab. 1. THE IDIOLECTS OF THE RUSSIAN SPEAKERS, after [Zimmerling, 2017].

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	F	M	F	F	M	F	F	F	F	M	M	F	F	F	F	F	M	M
	25	51	37	61	63	61	32	59	35	25	27	56	25	34	28	22	55	20
Main list	190	266	205	282	268	288	210	169	260	244	292	210	205	236	275	139	306	313
Add. list	5	0	9	14	9	20	6	7	12	10	25	6	10	6	41	1	27	38
	55,5%	77,7%	59,9%	82,4%	78,4%	84,2%	61,2%	49,4%	76%	71,3%	85,4%	61,4%	59,9%	69%	80,4%	40,6%	89,5%	91,5%

Tab. 2. THE IDIOLECTS OF THE BULGARIAN SPEAKERS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	F22	F21	F21	F22	F21	F59	F21	F22	F21	F21	M57	F47	F36	M41	F52	F58	M65	F16	F22
320	107	142	180	154	220	212	215	186	215	148	222	179	186	181	164	152	95	202	171
%	33,4%	44,4%	56,3%	48,1%	68,8%	66,3%	67,2%	58,1%	67,2%	46,3%	69,4%	55,9%	58,1%	56,6%	51,3%	47,5%	29,7%	63,1%	53,4%

Tab 3. Frequency classes and the ranking of DPS elements in BNC

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
m-measure	> 1000	400-1000	200-399	100-199	50-99	20-49	10-19	5-9	2-4	1	0
Total: 320	21	18	9	16	14	23	15	24	32	30	118

05.03.2017) and BNC (ca. 5,4 billion word forms, checked 23.04.2018)². We applied the ‘search-by-exact-form’ function, with filtering of the sample by annotators. The search has been narrowed through the approximation measure, so called *m*-measure (*m*) showing the use of the DPS elements in one dedicated context: contact position (distance [-1; 1]) of the subject dative pronoun in 1Sg to the predicative. The approximation is justified since most DPS are oriented towards the use in the 1Sg present indicative. The filters applied for processing Bulgarian data are discussed in (Ivanova, 2018).

The DPS elements were divided into 11 frequency classes. The highest rank ‘I’ was assigned to items with *m* > 1000, while the lowest rank ‘XI’ was assigned to items with *m* = 0.

Those DPS items which have a high frequency in BNC, almost always have high values of *Socio*. 20 from 21 elements in the frequency class I have the approval rate 1³ *Socio* 3⁴. This group includes -o-forms *приятно, трудно, известно, ясно, интересно, студено, достатъчно, лесно*, the nominal predicatives *жал, страх* and the set phrase *все едно*.

It is noteworthy that 37% of the test set of DPS stimuli (118 items) is not attested in BNC at all. One part of these items has low values of *Socio*, which indicates that such DPS as *ръмливо ми е* lit. ‘to-me is raining’, *навъсено ми е, правилно ми е, уместно ми е, симетрично ми е* etc. are neither considered standard nor are wide-spread. The other part of the low frequency DPS items is lacking or underrepresented in BNC because of their colloquial flavor. E.g. Bulg. *спешино* has *Socio* = 3, but only 5 occurrences in our BNC sample, Bulg. *мързеливо и никакво* have *Socio* = 6 and 0 occurrences in our BNC sample. These mismatches are due to the fact that BNC has only 1% of oral texts, while the DPS construction is oriented towards the speaker, hence the low frequency of many DPS items in the sample retrieved through the *m*-measure.

The sociolinguistic experiment shows that Bulgarian speakers are in general tolerant to DPS items with participial morphology, while BNC underrepresents them. E.g. *притеснено, напрегнато, замаяно, изморено, скапано, отпаднало* have high values of *Socio* (1 £ *Socio* £ 4), while in BNC *притеснено* и *напрегнато* have mid-range values (19 £ *m* £ 49). The remaining participial DPS items are hardly attested in BNC. Accusative predicatives show uneven distribution in both samples. The predicative *страх* has the highest *m*-value in BNC. The predicatives *грижа, срам, яд, гнус* have a high rank in both samples. *Еня* and *гъдел* have mid-range *m*-values, while *гняв, грях, мързел, студ* and *смях* are low frequency elements.

The frequency classes of Russian DPS predicatives in RNC are discussed in (Zimmerling, 2017).

Conclusions

The DPS construction in Russian and Bulgarian has a nucleus {Lex_k} which belongs to the lexicon and an extension which belongs to grammar {Gram_k}. The

methods of the double sociolinguistic and corpus analysis of the language-internal variation in the class of DPS predicatives proved operational for Bulgarian. The sociolinguistic experiment provided similar results for both languages, while the results of the corpus analysis reflect the differences in the size and structure of the chosen corpora³.

NOTES

1. See the web-page of the thematic block: <http://mpgu.su/isli/proektyi/tematicheskii-blok-imennyye-predikativyi-i-dativnyie-modeli_predlozheniya-v-slavyanskiih-yazyikah-na-mezhdunarodnom-sezde-slavistov-2018/>.
2. RNC: <http://www.ruscorpora.ru/>. BNC: <http://search.dcl.bas.bg/>
3. The paper was supported by the Russian Foundation for Basic Research, project № 18-512-18003.

REFERENCES

- Gradinarova, A.A. (2010). Impersonal constructions with dative subject and predicative ending in -o in Russian and Bulgarian. *Bolgarskaya rusistika*, 3/4, 34 – 55.
- Gradinarova, A.A. (2017). *Essays on Comparative Syntax of Bulgarian and Russian Languages*. Sofia: Iztok-Zapad.
- Gradinarova, A.A. (2018). Dative pattern with -o-predicative in Russian and Bulgarian languages at the background of its paradigmatic ties. *Russian language abroad*, 5, 18 – 23.
- Ivanova, E. Yu. (2018). Dative-predicative structures in Bulgarian: statistics of the experiment. *Russian language abroad*, 5, 11 – 17.
- Kulinich, E. (2018). Dative subject and stative predicates in Ukrainian. *Russian language abroad*, 5, 36 – 42.
- Maric, B. & Kerkez, D. (2018). Syntactic constructions of *the teško mi je and muka mi je* pattern in Serbian. *Russian language abroad*, 5, 43 – 50 [Марич, Б. & Керкез, Д. (2018).
- Mitkovska, L. (2018). The semantics of the Macedonian dative-predicative structures. *Russian language abroad*, 5, 24 – 29.
- Petrova, G. (2018). Nominal predicatives selecting dative experiencer in Bulgarian language: semantics and syntax. *Russian language abroad*, 5, 30 – 35.
- Zimmerling, A.V. (2010). Nominal predicatives and dative sentences in European languages (pp. 549 – 558). In: *Computational linguistics and intellectual technologies*, 9 (16): *Proceedings of the international conference "Dialogue 2010"*. Moscow: Russian State Univ. for the Humanities.

- Zimmerling, A.V. (2017). Russian predicatives in the perspective of experiment and corpus grammar (pp. 466 – 482). In: *Computational linguistics and intellectual technologies, 16 (23). T. 2.: Proceedings of the international conference "Dialogue 2010"*. Moscow: Russian State Univ. for the Humanities.
- Zimmerling, A.V. (2018a). Impersonal constructions and dative-predicative structures in Russian. *Voprosy Jazykoznanija*, 5, 7 – 33 [Циммерлинг, А.В. (2018а).
- Zimmerling, A.V. (2018b). Predicatives and predicates of state in Russian. *Slavistična revija*, 1, 45 – 64.
- Uhlik, M. (2018). Some features of Dative-predicative constructions in Slovenian. *Russian language abroad*, 5, 51 – 56.

✉ **Prof. Ivanova Elena Yu., DSc.**

St. Petersburg State University
7/9 Universitetskaya nab.
St. Petersburg, Russia
E-mail: e.y.ivanova@spbu.ru

✉ **Zimmerling Anton V., DSc.**

Pushkin State Russian Language Institute
6, Acc. Volgin St. (ulitsa Akademika Volgina, 6)

Moscow State Pedagogical University
Institute of Linguistics
Moscow, Russia

Russian Academy of Sciences
Moscow, Russia
E-mail: fagraey64@hotmail.com