Secondary Prepositions with Causal Meaning in Russian

Victor Zakharov[1], Kirill Boyarsky[2], Eugeny Kanevsky[3], and Anastasia Kozlova[1]

[1] Saint Petersburg University
Universitetskaya emb. 7-9
199034 Saint Petersburg, Russia
v.zakharov@spbu.ru, stasia.kozlova@gmail.com

[2] ITMO University
Kronverkskiy av. 49
197101 Saint Petersburg, Russia
boyarin9@yandex.ru

[3] Institute of Regional Economics Problems RAS
Serpukhovskaya St. 38
190013 Saint Petersburg, Russia
eak300@mail.ru

Abstract. The present study deals with Russian secondary prepositions, primarily focusing on multiwords. Secondary prepositions are units motivated by content words (nouns, adverbs, verbs), which may be combined with primary prepositions to form multiword prepositions (MWPs). Multiword prepositions perform the grammatical function of a preposition in a certain position of a syntactic structure in some contexts and can be a free word combination in others. This paper is devoted to analysis of the use of secondary multiword prepositions with causal meaning. We analyze the repertoire of Russian MWP causal prepositions and describe their statistical representation in corpora.

Keywords: Russian language, secondary prepositions, multiword prepositions, causal meaning, corpus statistics, parsing

1 Introduction

This study is part of a project whose goal is to create the first corpus semantic-grammatical description of Russian prepositional constructions. The Russian linguistic tradition implies the division of the class of prepositions into primary and secondary in origin, as well as into simple (single-word) and complex (multiword) units in structure. While the subclass of primary prepositions is relatively well studied and well documented, secondary prepositions have not received the same attention in the linguistic literature, despite making up the majority of prepositions as a class. The preposition is a common part of speech in many languages. It has been established that prepositions in Russian make up, on average, 10% of all tokens in each text [1].

The vague and complex semantics of prepositions underlies much of the debate about the nature of prepositions. Primary prepositions are very ambiguous. For example, the Russian preposition *в* 'in' has 23 meanings in the Dictionary of the Russian Language [2]. In other cases, the primary preposition is part of the secondary preposition. In total, there are several hundred secondary prepositions in the Russian language. Most often they can be considered as synonyms for primary ones.

We understand a prepositional meaning as a relation that occurs in prepositional constructions, where it should be considered as a special type of relationship between meaningful words. We consider this concept as a semigrammatical component of the language, linking fuzzy lexico-semantic classes of words. These relationships are established by a combination of a specific preposition, the semantic type of the lexeme that attaches the prepositional construction, and the case and the semantic class of the dependent word. Therefore, to describe and analyse causal prepositions, we provide parsing and identify the semantic classes of dependent words.

2   Related Work

Perhaps the most well-structured inventory of Russian secondary prepositions can be found in the Russian Grammar [3]. It is noted by the author(s) that a lot of the units listed are entities of uncertain part-of-speech status due to their preserved ability to include determiners and combine selectively with the other parts of the potential prepositional phrase [3, §1661]. The Explanatory Dictionary of Functional Parts of Speech of the Russian Language [4] contains less than 300 secondary prepositions. Much fewer, just 157, are found in the Explanatory Dictionary of Combinations Equivalent to a Word [5].

When speaking about causal relationship, a lot of studies are devoted to this type of relation ([6,7,8] and many others) and only several studies deal with causal prepositions [9,10,11]. However, these studies are not based on corpora. Our study relies on statistics on the use of these prepositions in large text material and on syntactic parsing.

3   Secondary Multiword Prepositions with Causal Meaning

Secondary prepositions are words and phrases that have assumed the function of a preposition. Structurally, these units can be divided into simple and complex (multiword) ones. Simple secondary prepositions are usually fully homonymous with some word form of their motivating content word or a different part of speech sharing the same root. The same words and word units may perform as prepositions as well as other parts of speech (e.g.: *силами* 'by force of' – preposition, noun, *снаружи* 'outside' – preposition, adverb, *исключая* 'excluding' – preposition, verb (participle).

Multiword prepositions (MWPs) make up a large part of secondary prepositions. Structurally speaking, a multiword preposition is a combination of a

content word and one or two simple adpositions. MWPs can be divided into nominal, adverbial or verbal units based on the part of speech of the motivating content word. Most MWPs contain only one adposition preceding or following the content word (e.g.: *рядом с* 'close to', *в результате* 'as a result'), but some include two adpositions enclosing the content element (e.g.: *в соответствии с* 'in accordance with', *по направлению к* 'toward, in the direction of'). The most commonly observed structural patterns of MWPs are Prep+N, Prep+N+Prep and Adv+Prep, where Prep stands for preposition, N for noun, Adv for adverb. Much like simple secondary prepositions, multiword prepositional units perform as prepositions in some contexts and as free word combination in others (e.g., preposition + noun: *в форме* 'in the form of', conjunction + preposition: *что до* 'as for'; verb + preposition: *начиная с* 'starting with').

As a rule, the distinction between these ambiguous entities is outlined neither in grammar books nor in dictionaries. The great variety of MWPs on all language levels implies the necessity of an in-depth analysis of their common features. In other words, we need to understand, firstly, what unites such diverse entities in order to be able to discern free combinations from MWPs.

As has already been stated, prepositional multiword entities do not always function unambiguously as MWPs. Although our current paper is devoted to causal prepositions, in order to define limits of all MWPs, we have formulated the following preliminary list of the main characteristic features of multiword prepositions:

– MWP performs the grammatical function of a preposition in a certain syntactic position as part of a prepositional phrase; that is, it governs a noun or a nominalised word (sometimes an infinitive).
– MWP inherits the semantics of the notion word (noun, verb); it derives from as well as its valency (*на основе* 'on the grounds of' – *основа чего?* 'the grounds of what?'; *в зависимости от* 'depending on' – *зависеть от чего?* 'to depend on what?'; *с целью* 'with the aim to' – *цель что сделать?* 'aim to do what?').
– As a rule, it contains one or two primary prepositions.
– Its nominal components tend to have abstract semantics.
– It has a relatively high frequency among multiword units of the same structural type.
– It is idiomatised, i.e., its nominal component loses its lexical meaning to an extent (which is why MWPs are sometimes called "prepositional idioms").
– The grammatical number of the noun cannot be changed (it is either singular or plural).
– It has a primary preposition as a synonym.
– In most cases, it does not allow for insertion or separation (as a rule, the noun cannot have a possessive or adjectival determiner).
– All of these features are characterised by significant statistical regularity.

4   Material

Causal preposition meaning is the meaning of constructions where the preposi-
tional group indicates the cause of an action or the influencing factor. The word
'cause' is the basic term used to interpret the whole lexical composition, which
is associated with the category of determining the cause. Dictionaries of a lan-
guage divide two meanings of this word: 1) cause as a phenomenon that in-
advertently causes another phenomenon, ontological cause; 2) cause as a basis,
precondition for the realization of an event, action, i.e., subjective, explainable
cause. The meaning of prepositions is based on that description. Among ways
to express causal relations, the most common are prepositional-case forms and
complex sentences with a subordinate causal part, and most of causal conjunc-
tions come from prepositions.

Causal relationships can be expressed in Russian by some primary preposi-
tions and a large number of secondary ones. They can enter into connection with
nouns and pronouns in genitive, dative, accusative and instrumental. The lists
of causal prepositions differ in different sources. According to our analysis, the
list is as follows: *за, из, из-за, на, от, по, под, после, при, с, через* (primary prepo-
sitions), *благодаря* 'thanks to', *в зависимости от* 'depending on', *в ответ на* 'in
response to', *в результате* 'as a result of', *в свете* 'in light of', *в связи с* 'due to', *в
силу* 'by force of', *за счёт* 'on account of', *исходя из* 'drawing from', *на основании*
'on the basis of', *на основе* 'based on', *на почве* 'on the ground of', *по причине* 'be-
cause of, for the reason of' (secondary prepositions). (Here we provide English
equivalents only for the secondary prepositions because the meanings of the pri-
mary ones are highly context-dependent). These prepositions form clusters of
synonymy. Different prepositions can express the same meanings and grammat-
ical relations when used in the same phrases. In the sentences *Он не пришел по
причине болезни – Он не пришел из-за болезни – Он не пришел вследствие болезни*
('He did not come because of the disease') it is possible to interpret the prepo-
sitions as synonyms.

Our selection consists of 13 multiword preposition candidates that have been
observed to express causal or causal-adjacent relations:

– *в зависимости от* 'depending on'
– *в ответ на* 'in response to'
– *в преддверии* 'on the eve of, at the forefront of'
– *в результате* 'as a result of'
– *в свете* 'in light of'
– *в связи □* 'due to'
– *в силу* 'by force of'
– *за счёт* 'on account of'
– *исходя из* 'drawing from'
– *на основании* 'on the basis of'
– *на основе* 'based on'
– *на почве* 'on the ground of'
– *по причине* 'because of, for the reason of'

The results of the statistical analysis presented in this article have been acquired on the Russian National Corpus (RNC, www.ruscorpora.ru). This corpus was chosen due to its considerable size (about 375 million tokens). For parsing and manual (intelligent) analysis, random samples of 500 sentences were taken from the RNC for each preposition.

5   Results

5.1   Structural Features of Causative MWPs

The first set of features to observe is the structure of the MWPs in question. 10 out of 13 units are bigrams of a content word and a simple adposition, which appears to be the most typical MWP structure. 9 of the bigram units follow the structural pattern of Prep+Noun, one has the less common pattern of Verb+Prep. The remaining 3 out of 13 units are trigrams consisting of a noun between two adpositions.

As has been noted by us in [12] the three simple adpositions most commonly used as elements of multiword prepositions are *в, на, по*. Out of the 13 items under current study, 7 contain the preposition *в*, 4 contain *на*, 1 contains *по*.

Most of the content words in the causative MWPs refer to the two nodes of causal relations: the reason (*основание* 'foundation', *основа* 'base', *почва* 'ground', *причина* 'reason') and the effect (*ответ* 'response', *результат* 'result'), as well as the relation itself (*зависимость* 'dependency', *связь* 'connection', *сила* 'force'). The semantics of the motivating content words corresponds with the observed tendency of MWP component nouns to lean towards abstraction.

5.2   Statistical Analysis of Causative MWPs

Another point of interest is the use (frequency) of the content (base) words as multiword unit (MWU) components in comparison to their general corpus frequency. The table below demonstrates the frequencies of the content words in question as well as the frequencies of the multiword unit themselves (Table 1).

As demonstrated by the table, most of the content words retain their relative independence as they are not bound to the other parts of the MWUs. Only two of them, *в преддверии* and *исходя из*, appears to be a set phrase in which the motivating word is not almost used without the corresponding preposition.

To gain a clearer understanding of how prepositional units function as MWPs or free combinations we have studied the use of the causative prepositional units in the corpus. The results are presented in Table 2.

The results show that these word combinations are mostly used as prepositions and not free word combinations. MWPs which demonstrate a relatively low percentage of prepositional sense have a strong sememe as the base word that retains its original meaning even as part of a preposition. For example, in the preposition *на основе* 'on the basis of', the base word 'base' tends to retain approximately the same meaning both in the preposition and outside it.

Table 1: Frequency counts of MWUs in relation to base word frequencies in RNC

Base word	Multiword unit (MWU)	MWU, count	Base word, count	MWU/base word, %
Исходя[4]	*Исходя из(о)*	5314	5788	91,8
Преддверие[5]	*В преддверии*	1134	1373	82,6
Зависимость	*В зависимости от(о)*	9859	21604	45,6
Результат	*В результате*	29745	93454	31,8
Счёт	*За счё/ет*	16113	52693	30,6
Связь	*В связи с(о)*	23496	96248	24,4
Основа	*На основе*	12764	52299	24,4
Основание	*На основании*	11296	48353	23,4
Почва	*На почве*	2929	25772	11,4
Сила	*В силу*	14884	266166	9,0
Ответ	*В ответ на*	6723	77869	8,6
Причина	*По причине*	4557	81422	5,6
Свет	*В свете*	5727	156587	3,7

Table 2: Percentage of prepositional use of MWP candidates in Russian National Corpus

MWP	% of prepositional use
Исходя из(о)	100.0
В результате	100,0
По причине	100,0
В зависимости от(о)	99.8
В связи с(о)	99.4
На основании	99.8
В ответ на	99.8
В преддверии	98.4
За счё/ет	96.2
На почве	96.6
В силу	83.6
На основе	79,0
В свете	49.2

For example, *на основе* was found to be a free word combination in contexts referring to the physical basis of an entity, e.g. [X] *"на основе гиалуроновой кислоты"* ('hyaluronic acid-based [X]'); similarly, word combination *в свете* 'in light of' was used literally in contexts where the governee belonged to the semantic class of objects capable of emanating light, e.g. *"в свете заходящего солнца"* ('in the light of the setting sun'). The MWP candidate *в силу* 'by force of, due to' was found to be used occasionally as a free combination (*[верить] в силу* '[believe] in the force') and as part of an adverbial idiom (*[вступить] в силу* 'come into power'), which led to the relatively lower observed percentage of its prepositional use as well.

A special point of interest is the separability of MWPs, that is, the allowance for modifier insertion into the MWP structure. To study this phenomenon, we

have examined context samples of some causative MWP candidates with and without content word modifiers. Overall, our presupposition that insertion is atypical for MWPs has been proven true. Since most of the content words in our MWP candidate selection are nouns, it was primarily adjectival modifiers that were found splitting the original prepositional unit structure. As it was said, the nominal component of MWPs loses often its lexical meaning. When modifier insertion takes place, the semantic weight of the whole construction figuratively shifts back to the modified noun, which retains its original lexical meaning. Therefore, the resulting structure can no longer perform the prepositional function and can only be regarded as a free word combination. Some nominal causative MWPs do not seem to lose their function in the case of anaphoric use of personal pronouns, such as *на его почве* 'on his [its] ground', *в её преддверии* 'on her [its] eve'.

The meaning of a secondary preposition depends sometimes on the meaning of semantics of a governing word and of a governee (dependent word). We have obtained governee lists (see Table 3) for each of the prepositional units in question by the SemSin parser [13], which builds a dependency tree for each sentence and detects types of relations between its nodes. The parser relies on a semantic-syntactic dictionary and a classifier, both of which are extensions of the semantic dictionary by V.A. Tuzov.

The most frequently occurring dependent words for each set phrase often quite clearly illustrate the meaning of construction, whether it is a preposition or not. It is also noticeable that for the preposition with the lowest percentage of prepositional meaning, *в свете* 'in the light of', one of the most frequent dependent words is 'campfire': it becomes obvious that this phrase is often used as free combination. Dependent word classes influence the fact of whether the word is used as a preposition component within the MWP: the more dynamic the meaning of the semantic class is, the more likely it is to be non-prepositional.

6   Conclusion

To conclude, it can be noted that analyzed MWUs quite often perform the function of prepositions (MWPs). We can conclude also that causative MWPs generally do not allow for insertion (modification of the content component) except when the modifier is a personal pronoun modifying the nominal component or a particle modifying the verbal component. Whether this rule applies to the entirety of the MWP class is subject to further investigation.

The most frequent governors in prepositional usage cases belong to the group of verbs and verbal nouns expressing change of state, e.g., *получить* 'receive', *возникать* 'appear', *образоваться* 'form', or expressing difference, e.g., *меняться* 'change', *варьироваться* 'vary', *отличаться* 'differ; for the preposition *в зависимости* □□ 'depending on'. For the preposition *в силу* 'due to, by force of' the governors were useful in identifying free usage cases, e.g., *вступление* 'entry', *вступить* 'come', *верить* 'believe' [in(to) (the) power]. Inversely, the homonymy resolution of the MWP candidates *в свете, в преддверии* was more

Table 3: Most frequent governees and governees semantic slasses in causal MWP constructions

MWP	The most frequent governees	The most frequent semantic classes
В преддверии	*выборов* 'elections'	*Обладание* 'Possession'
	года 'of the year'	*Конкретное_время* 'Specific_time'
	юбилея 'anniversary'	*Занятие* 'Occupation'
В зависимости от(о)	*того* 'that'	*Спецкласс_для_ТО_и_ЭТО* 'Special_class_for ТО_and_ЭТО'
	условий 'conditions'	*Событие* 'Event'
	типа 'like'	*Качество* 'Quality'
Исходя из(о)	*этого* 'that'	*Сообщение* 'Message'
	того 'that'	*Этот-Другой* 'This_Other'
	опыта 'experience'	*Дух* 'Spirit'
За счё/ет	*средств* 'means'	*Деньги* 'Money'
	использования 'of using'	*Изменение* 'Change'
	того 'that'	*Занятие* 'Occupation'
В связи с(о)	*этим* 'by this'	*Этот-Другой* 'This_Other'
	развитием 'development'	*Событие* 'Event'
	делом 'case'	*Дело* 'Business'
На основе	*анализа* 'analyses'	*Науки* 'Science'
	данных 'data'	*Сообщение* 'Message'
	опыта 'experience'	*Документы* 'Documents'
На основании	*данных* 'data'	*Сообщение* 'Message'
	того 'that'	*Нечто* 'Something'
	анализа 'analysis'	*Документы* 'Documents'
В результате	*которой* 'which'	*Изменение* 'Change'
	работы 'work'	*Сообщение* 'Message'
	деятельности 'activity	*Нечто* 'Something'
По причине	*того* 'that'	*Событие* 'Event'
	отсутствия 'absence'	*Дух* 'Spirit'
	болезни 'illness'	*Душа* 'Soul'
В ответ на	*вопрос* 'question'	*Сообщение* 'Message'
	это 'this'	*Этот-Другой* 'This_Other'
	просьбу 'request'	*Душа* 'Soul'
На почве	*любви* 'love'	*Душа* 'Soul'
	ревности 'jealousy'	*Дух* 'Spirit'
	отношений 'relations'	*Борьба* 'Struggle'
В свете	*событий* 'events'	*Событие* 'Event'
	сказанного 'said'	*Сообщение* 'Message'
	костра 'campfire'	*Стихия* '□lement'
В силу	*причин* 'reasons'	*Событие* 'Event'
	того 'that'	*Причина* 'Cause'
	обстоятельств 'circumstances'	*Дух* 'Spirit'

successful in the presence of their governees. As such, contexts with the governess *фары* 'headlights', *фонари* 'streetlamps', *луна* 'the moon' for the preposi-

tion *в свете* 'in light of' and *рот* 'mouth', *влагалище* 'vagina' for the preposition *в преддверии* 'at the forefront of' were found to be free word combinations. Another interesting observation is that high number of the causative prepositional units were found to take the initial position in a sentence, or a clause as evidenced by the inclusion of punctuation marks and conjunctions.

In the future, we plan to describe MWPs with other meanings.

References

1. Lyashevskaya, O.N., Sharov, S.A.: Frequency Dictionary of the Modern Russian Language (on the Materials of the National Corpus of the Russian Language). Azbukovnik, Moscow (2009).
2. Dictionary of the Russian Language, 4 volumes. 3rd edn. Russkiy yazyk, Moscow (1988).
3. Shvedova, N.Ju.: Russian Grammar. Vol. 1: Phonetics. Phonology. Word Stress. Intonation. Word Formation. Morphology. (in Rus). Nauka, Moscow (1980).
4. Efremova, T.F.: Explanatory dictionary of functional parts of speech of the Russian language. AST, Moscow (2004).
5. Rogozhnikova, R.P.: Explanatory dictionary of combinations equivalent to a word: Approx. 1500 fixed phrases of the Russian language. AST, Moscow (2003).
6. Diessel, H., Hetterle, K.: Causal clauses: A crosslinguistic investigation of their structure, meaning, and use. In: P. Siemund (ed.). Linguistic Universals and Language Variation. Berlin, Boston: De Gruyter Mouton, p. 21 – 52 (2011).
7. Kroupová, Libuše. Vztah významu gramatického a lexikálního u předložek. In: Slovo a slovesnost, 1980, roč. 41, č. 1, s. 49 – 52 (1980).
8. Vsevolodova, Maya – Yashchenko, Tatyana: Causal relationships in modern Russian. Moscow: Russkiy yazyk. 208 p. (in Russian) (1988).
9. Levontina, Irina: <Causal prepositions>. In: New explanatory dictionary of synonyms of the Russian language. Moscow, p. 144 – 152. (in Russian) (1997).
10. Iordanskaya, Lidiya – Melchuk, Igor: On the semantics of Russian causal prepositions. In: Linguistic Journal. No. 2, p. 162–211. (in Russian) (1996).
11. Luraghi, Silvia. Prepositions in Cause expressions. In: Papers on grammar, vol. 12, No. 2, p. 609–619 (2005).
12. Zakharov, V., Azarova, I.: Semantic structure of Russian prepositional constructions. In: K. Ekstein, V. Matousek (eds.). Lecture Notes in Computer Science, vol. 11697 LNAI (Text, Speech, and Dialogue - 22th International Conference, TSD 2019 Proceedings), pp. 224-235. Springer International Publishing AG (2019).
13. Boyarsky K., Kanevsky Ye. Semantic-syntactic parser SemSin [Semantiko-sintaksicheskiy parser SemSin]. In: Scientific and technical bulletin of information technologies, mechanics and optics [Nauchno-tekhnicheskiy vestnik informacionnykh tekhnologiy, mekhaniki I optiki]. Vol.15. No. 5, pp. 869-876 (2015).