

Synonymy in Modern Tatar reflected by the Tatar-Russian Socio-Political Thesaurus

Alfia Galieva

Tatarstan Academy of Sciences

E-mail: amgalieva@gmail.com

Abstract

This paper discusses some aspects of lexical synonymy in the modern Tatar language that have been revealed in the course of compiling the bilingual *Russian-Tatar Socio-Political Thesaurus*. Building the thesaurus is aimed at fixing all Tatar single words and multiword items related to the socio-political sphere with their Russian equivalents. A distinguishing feature of the contemporary Tatar lexicon is a great deal of absolute synonyms which emerged due to a combination of intralinguistic and extralinguistic factors. We disclose social and linguistic causes of the emergence of synonyms, describe the main structural types of synonymous items, and present corpus data on their frequency. Corpus data prove that synonymy in socio-political terminology is rather an artificial and superficial phenomenon. Currently most Tatar socio-political terms are coined by calquing the corresponding Russian terms, and lexical preferences of translators and terminology developers may differ, which leads to a large number of competing items of different origin and structure. On the level of multiword items, lexical variation is complicated by the factor of syntactic variation, which in its turn multiplies the number of synonymous compounds. Parallel denominations are used for a wide range of phenomena, including official names of state structures and social institutions.

Keywords: lexical synonymy, absolute synonyms, socio-political vocabulary, bilingual thesaurus, the Tatar language

1 Introduction

Collecting lexical data, systemizing it and mapping semantic relations between the lexical items are the important stages of lexicographic work, ones that gives material for comprehending current processes in languages. In this paper, we discuss some facts of lexical synonymy in the modern Tatar language that have been encountered in the process of compiling the bilingual *Russian-Tatar Socio-Political Thesaurus*. The Tatar language of recent decades has experienced a stage of renewal which caused significant changes in vocabulary, including the emergence of a large number of synonymous items. In present work, the main attention is paid to absolute synonymy.

The conventional viewpoint is that true synonyms are rarely found in a language. However, the situation with Tatar socio-political vocabulary in many respects contradicts this thesis about the rareness of absolute synonymy; this phenomenon is observed due to a set of extralinguistic and intralinguistic causes, which require explanation.

The body of the paper is organized as follows. Section 2 gives a brief description of the current language situation in Tatarstan, and puts forwards an idea about the social motivation of changes in Tatar vocabulary. Section 3 presents the main goals and methodology of compiling the *Russian-Tatar Socio-Political Thesaurus*. Section 4 outlines the basic theoretical background of the study. Section 5 describes the most important aspects of lexical synonymy in the modern Tatar language that have been registered in the course of compiling the bilingual thesaurus; synonymy at the level of single

words and multiword items is discussed; corpus evidence about the distribution of synonyms is given. Section 6 lists the conclusions and outlines the prospects for future work.

2 Social Motivation of Changes in Tatar Vocabulary

An important feature of the contemporary Tatar language is the emergence of a large number of synonymous items, this phenomenon being caused by a set of intralinguistic and social factors, which requires an explanation.

Tatarstan is a republic of the Russian Federation which is located in the Volga region of the European part of Russia, with the capital city of Kazan. According to the 2010 census, the population of Tatarstan consists of 53% of ethnic Tatars, 40% of ethnic Russians, 7% of people of other ethnic origin. At present practically all Tatars speak Russian (but not vice versa).

By the late 1980s, the minority languages of the USSR had been largely supplanted by Russian in the urban areas of most Soviet republics; ethnic groups used Russian in the public sphere, while state and public administration and office work were also conducted in Russian. The danger of losing the national language was one of the main driving forces of the ethnic renaissance that developed in Tatarstan in the late 1980s, as well as in other ethnic regions of the USSR.

The legal basis for promoting the use of the Tatar language in the public sphere emerged with the adoption of the Language Law in 1992, which triggered the implementation of the government-sponsored Tatar language revival program and the establishment by state and municipal bodies of special committees for Tatar language use. Active translation work was thus launched in state structures and institutions, and an urgent need in Tatar socio-political terminology arose.

Currently in accordance with the Constitution of the Republic of Tatarstan and the Language Law, the two state languages of the republic are Tatar and Russian, which formally have equal legal status. In particular, the texts of laws of the Republic of Tatarstan and other legal acts are to be published in both Tatar and Russian, while state authorities and institutions use both state languages of the Republic of Tatarstan.

So the late 1980s had an active social impact on the language situation in Tatarstan, and the Language Law significantly raised the status of the Tatar language. At the same time a religious revival began in Tatarstan; Tatars recognized themselves as a part of the vast Arab-Muslim cultural area, and direct cultural connections with the Arabic countries and Turkey were initiated.

The late 1980s and 1990s thus became a period of intensive linguistic innovation in Tatarstan. A new regard for Tatar culture and the increased social role of the Tatar language generated a movement for purifying it from numerous Russian and Western international words. As a result, the available vocabulary underwent revision; numerous loanwords of Russian and European origin were offered to be substituted with their Tatar, Arabic or Persian equivalents; many Arabic and Persian words that in the Soviet period were regarded as obsolescent, returned to the active lexicon again. Consequently a large number of new words appeared to designate existing and newly created realities, and Tatar vocabulary was enriched by numerous synonymous items.

So the vast changes in Tatar vocabulary of the last decades were socially motivated. The location of Tatar culture at the intersection of Occidental and Oriental civilizations leads to active language contacts both with the Arab-Muslim and the European cultural areas. At the same time, the dominant role of Russian as the state language of the Russian Federation remains the main cause of a huge number of words and constructions calqued (translated component-by-component) from Russian.

Currently most Tatar socio-political terms are coined by calquing the corresponding Russian terms. Diverse groups of specialists with dissimilar world-views and ideological guidelines develop terminology in Tatar; in terminology design they may principally be oriented towards different cultural spaces with different languages spoken:

- Tatar and Turkic vocabulary;
- Arabic and Persian vocabulary;
- international Greek, Latin or English vocabulary;
- Russian vocabulary, etc.

The heterogeneous preferences of terminology developers as well as of other language users lead to proposing and choosing different designations for the same entity. In these circumstances there is a need for new lexicographic resources that would embrace synonymous items and cover all existing variants of terms to provide users with more effective access to information.

3 *Tatar-Russian Socio-Political Thesaurus: Main Goals and Methodology of Compiling the Work*

For low-resource languages like Tatar, compiling special domain vocabulary is a challenge in many respects:

- insufficient degree of terminology development that leads to the absence of relevant designations for special language concepts
- a lack of special sources for terminology compiling;
- a lack of terminology management, etc.

The project of developing the *Russian-Tatar Socio-Political Thesaurus* (<http://tattez.antat.ru/>) is aimed at compiling the whole body of modern Tatar vocabulary related to the following basic domains: state government, economy, social life, justice, warfare, culture, religion. The thesaurus also comprises some general lexicon branches representing lexical items which can be found in various domain specific texts. All items have Russian equivalents (correspondences) and are represented at concept and lexical entries levels which are arranged hierarchically.

This new lexical resource is being developed on the basis of the Russian RuThes (http://www.labinform.ru/pub/ruthes/index_eng.htm) thesaurus. Both thesauri are implemented as a hierarchy of concepts viewed as units of thinking. Each concept is linked with a set of language expressions (single words and multiword expressions) that refer to it in texts (lexical entries). Each RuThes concept is a set of synonyms or near-synonyms (plesionyms). RuThes developers, Loukachevich and Dobrov, use a weaker term, *ontological synonyms*, to designate words belonging to different parts of speech (like *stabilisation, to stabilise*) and related to different styles and genres; idioms and even free multiword expressions which are potentially synonymous to single words are also included (Loukachevich & Dobrov 2014).

The conceptual structure of RuThes determines the direction of concept development in the Tatar part, and the basic structure of the conceptual relations of RuThes is preserved, i.e. the Tatar component is based upon the list of RuThes concepts.

The general methodology of creating the Tatar part of the thesaurus includes the following steps.

1. Search for equivalents (corresponding words) which are actually used in Tatar as translations of Russian words.

2. Adding new concepts representing topics which are important for the socio-cultural life of the Tatar society and are not presented in the original RuThes (the list of required vocabulary is compiled, the concept names and lexical entries are distinguished and arranged according to RuThes structure).
3. Revising relations between the concepts considering the place of each new concept in the hierarchy of the existing ones and, if necessary, adding the new concepts of the intermediate level (Galieva et al. 2017).

The Thesaurus is mainly being compiled by manual translation of terms from the Russian RuThes into Tatar. Tatar language specific concepts and their lexical entries are also added, so each part of the Thesaurus – the Russian and Tatar ones – represents a unique language-internal system of lexicalizations. At the same time, the languages are interconnected so that it is possible to go from the concepts and words in one language to the corresponding items in the other.

The main challenge of working on this project is concerned with acquiring lexical data and representing Tatar socio-political vocabulary as fully as possible, including a large number of synonymous items in actual use. Searching for translation equivalents and adequate correspondences that are practically used in texts becomes, in many cases, a laborious and time-consuming task. The available bilingual Russian-Tatar dictionaries of general lexicon and special dictionaries are outdated, and even new lexicons do not contain the required items or include obsolete lexical material (See (Galieva et al. 2017)).

In the process of compiling the thesaurus, data from the following available Tatar corpora is used:

1. Tatar National Corpus (<http://tugantel.tatar/?lang=en>);
2. Corpus of Written Tatar (<http://www.corpus.tatar/en>).

These corpora include texts of various genres, from official documents and scientific publications to media texts, fiction, and textbooks. They are being permanently replenished, which provides a constant inflow of fresh linguistic material. The corpora have comparable volumes, each containing more than 100 million tokens, and are supplied with a system of morphological annotation (Suleymanov et al. 2013; Nevzorova et al. 2015). The data provided by these corpora allow us to acquire reliable information on meanings, typical contextual relations and frequency of use of Tatar words, which is a necessary stage in compiling a thesaurus. Currently the *Russian-Tatar Socio-Political Thesaurus* contains 8,000 concepts provided with lexical entries.

4 Theoretical and Lexicographic Aspects of Synonymy: Brief Outline

Synonymy is one of the fundamental concepts in linguistics; it manifests itself at different levels of the language. A prevailing point of view in linguistics is that lexical synonymy shows a high degree of development of a language, enabling us to encode the most subtle differences between entities or ideas and different perspectives of viewing them.

The notion of synonymy has been subject to a large number of studies as to what it constitutes, where its borders lie and what its scale is, and there are a number of controversial interpretations. Researchers distinguish between cognitive synonyms (Cruse 1986) and near-synonyms, or plesionyms (Hirst 1995; Edmonds & Hirst 2002; Divijac 2006; 2010, Desagulier 2014).

Mapping synonyms in special lexicographic resources demands some kind of technical approach, because the developers cannot carry out a meticulous study on semantic and contextual differences of thousands of words represented in a resource. A synonym, in *Webster's New Dictionary of Synonyms*,

means “one of two or more words in the English language which have the same or very nearly the same essential meaning” (Egan 1984: 24 a).

Synonymy is the basis for organizing lexical items in Princeton WordNet (Miller et al. 1990, Miller 1995) and other WordNet-like resources (Vossen 2002). Miller and Fellbaum suggest a vaguer term for synonymy, namely “semantic similarity”, which is defined as follows: “two expressions are synonymous in a linguistic context C if the substitution of one for the other in C does not alter the truth value” (Miller et al. 1990).

In *EuroWordNet General Document* words are considered to be synonyms if they “denote the same range of entities, irrespective of the morpho-syntactic differences, differences in register, style or dialect or differences in pragmatic use of the words” (Vossen 2002: 18). Moreover, developers rely upon another, more practical, criterion which follows from the homogeneity principle that synonyms cannot have any other semantic relation (Vossen 2002: 18).

The notion of absolute synonymy also leaves room for different interpretations. Cruse maintains that two lexical units would be absolute synonyms if and only if all their contextual relations were identical (Cruse 1986: 168).

According to Lyons, two (or more) synonymous expressions are absolute synonyms if they satisfy the following three conditions:

- 1) all their meanings are identical;
- 2) they are synonymous in all contexts;
- 3) they are semantically equivalent (i.e. their meaning or meanings are identical) in all the dimensions of meaning, both the descriptive and the non-descriptive ones (Lyons 1995: 61).

The attitudes in related works range from total rejection of absolute synonymy in a language to claiming their extreme uncommonness (Ulmann 1962; Lyons 1995; Cruse 1986 and others). Cruse emphasizes that “natural languages abhor absolute synonyms just as nature abhors a vacuum” (Cruse 1986: 270), and maintains that it would be reasonable to assume “that if the relationship were to occur, it would be unstable”, because one of the items would fall into obsolescence, or some differences between the items would develop (Cruse 1986: 270). Synonymous items are usually distinguished from each other by a set of individual senses, added implications and connotations, or idiomatic use.

5 Synonymy in Current Tatar Socio-Political Vocabulary

5.1 Synonymy of single words

In the Soviet era many words denoting socio-political realities were borrowed from Russian and European languages or were coined by component-by-component translation of Russian items. A movement for the Tatar language revival led to an active language renewal, which enriched the Tatar vocabulary with a great number of synonymous items of Tatar and Oriental (Arabic and Persian) origin. Many Arabic and Persian words and their derivatives considered as obsolescent in the Soviet period (and correspondingly marked in dictionaries of the time) were thus entered into the active vocabulary fund. So in modern Tatar, words of different roots (Turkic, Russian, Arabic, Persian, Greek, Latin, and English) denoting the same referent, are competing. Selection of homosemous words is not completed yet, and the concurrence brought dissimilar results for different lexemes. Corpus data allows us to observe and fix the process of coining new terms, although this process is laborious and contradictory in many aspects. Table 1 represents one-word synonyms of European and Russian origin which are currently being displaced with equivalents of Oriental (Arabic and Persian) and Tatar

(Turkic) origin. It should be noted that words of Oriental origin in Table 1 abound in present-day official and media texts which underwent preliminary editorial revision. However, in actual oral and especially colloquial speech, words of Russian origin and international words and their derivatives are actively used.

Table 1: Examples of words of European and Russian origin which are presently being displaced with Oriental and Turkic equivalents.

Words	Word origin	English translation	Frequency in Corpus of Written Tatar / Tatar National Corpus	Type of texts of words using
<i>экономик</i>	Greek	'economic'	621 / 676	Soviet official and media texts
<i>икътисади</i>	Arabic	'economic'	27,351 / 22,274	Present-day official and media texts
<i>политик</i>	Greek	'political'	1,005 / 1,536	Soviet official and media texts
<i>сәяси</i>	Arabic	'political'	25,011 / 24,489	Present-day official and media texts
<i>страхование</i>	Russian	'insurance'	239 / 109	Soviet official and media texts
<i>страховкалау</i>	Tatar	'insurance'	814 / 1,331	Soviet official and media texts
<i>иминиятләштерү</i>	Tatar	'insurance'	2,222 / 1,221	Present-day official and media texts
<i>иминият</i>	Arabic	'insurance'	3,894 / 1,667	Present-day official and media texts
<i>больница</i>	Russian	'hospital'	2,832 / 6,383	Soviet official and media texts
<i>хастаханә</i>	Persian	'hospital'	19,675 / 13,397	Present-day official and media texts
<i>сырхауханә</i>	Persian	'hospital'	3,492 / 1,235	Present-day official and media texts
<i>налог</i>	Russian	'tax'	547 / 1,423	Soviet official and media texts
<i>салым</i>	Tatar	'tax'	23,223 / 17,322	Present-day official and media texts

Loanwords produced stems for sets of derivatives of different structure, which compete in their turn (like the Tatar words *страховкалау* and *иминиятләштерү* derived from corresponding Russian and Arabic stems, represented in Table 1).

Words of Oriental and Turkic origin do not always dominate; in many cases international and Russian words demonstrate high frequency in written texts. Table 2 presents lexical items which retained their positions despite the emergence of Oriental competitor words.

Table 2: Examples of international and Russian loanwords maintaining high frequency

Words	Word origin	English translation	Frequency in Corpus of Written Tatar / Tatar National Corpus
<i>республика</i>	Latin	'republic'	316,667 / 258,433
<i>жәмһүрият</i>	Arabic		2,479 / 1,631
<i>суд</i>	Russian	'court of law'	15,725 / 13,873
<i>мәхкәмә</i>	Arabic		3,203 / 4,957
<i>компьютер</i>	English	'computer'	13,469 / 10,032
<i>санак</i>	Turkic		631 / 67

Words of Tatar (Turkic) origin that are formed of different stems or of the same stem may also compete. Notable examples are Tatar words referring to *businessman* and *business* (entrepreneurship), from the same Turkic stem *эш* 'affair, business'. Table 3 presents the quantitative distribution of synonymous words denoting *businessman*, in corpus collections, and the number of their most frequent derivatives.

Table 3: Native Tatar words with a business-related meaning and their most frequently used derivatives

Lexeme	English translation	Frequency in Corpus of Written Tatar	Frequency in Tatar National Corpus
<i>эшмәкәр</i>	‘businessman’	22,671	13,839
<i>эшкуар</i>	‘businessman’	8,606	11,289
<i>эшмәкәрлек</i>	‘business, entrepreneurship’	6,454	4,209
<i>эшкуарлык</i>	‘business, entrepreneurship’	2,641	3,859

To assess the semantic similarity of words we may consider their distributional similarity using corpus data. It is easy to see that words *эшмәкәр* and *эшкуар* are actively used in present-day socio-political texts; they have the same meaning and are characterized by essentially identical contextual environment (see Table 4). So we may conclude that words *эшмәкәр* and *эшкуар* are absolute synonyms. The same can be said about other examples in the tables above.

Table 4: The most frequent collocations of words *эшкуарлык* and *эшмәкәрлек*

Typical collocations of word <i>эшкуарлык</i>	Frequency in Corpus of Written Tatar	Typical collocations of word <i>эшмәкәрлек</i>	Frequency in Corpus of Written Tatar
<i>эшкуарлык комитеты</i> ‘business committee’	341	<i>эшмәкәрлек субъекты</i> ‘business entity’	447
<i>эшкуарлык субъекты</i> ‘business entity’	186	<i>эшмәкәрлек комитеты</i> ‘business committee’	183
<i>эшкуарлык эшчәнлеге</i> ‘business activity’	94	<i>эшмәкәрлек үсеше</i> ‘development of business’	269
<i>эшкуарлык үсеше</i> ‘development of business’	65	<i>эшмәкәрлек эшчәнлеге</i> ‘business activity’	257

In many cases absolute synonyms form nests of elements of different frequency. In Table 5 one can find the distribution of words referring to *patriotism*.

Table 5: Synonyms of different origin and structure denoting *patriotism*.

Words	Word’s origin	Frequency in Corpus of Written Tatar	Frequency in Tatar National Corpus
<i>патриотизм</i>	Greek	763	703
<i>патриотлык</i>	Tatar derivative from the Greek stem	446	105
<i>ватанпәрвәрлек</i>	Tatar derivative from the Arabic stem	853	797
<i>ватандарлык</i>	Tatar derivative from the Arabic stem	80	83
<i>ватанчылык</i>	Tatar derivative from the Arabic stem	12	30

To declare that the words above are absolute synonyms, we proceed from the assumption (which can be debated) that we do not observe a real word borrowing here (new or unknown referents with their unique links do not appear), rather one verbal label is changed by another. For example, international *economics* was arbitrarily replaced by Arabic *икътисад*, consistently and in all collocations, and ordinary Tatar speakers may not know anything about the real Arabic word denoting economics, its senses, collocations and connotations in the Arabic language. So we observe a rather superficial change of lexical items with one another.

5.2 Synonymy of Multiword Expressions and Compound Terms

On the level of multiword terms and phrases synonymy is complicated by grammatical factors. An interesting issue concerns the absence of primordial relative adjectives in Turkic languages. A great number of Russian terminological word combinations contain relative adjectives. Because of a comparatively small number of relative adjectives in Tatar (all of them borrowed from Oriental and European languages or from Russian), when translating such multiword items the so called *ezāfe* constructions are usually formed by two nouns. *Ezāfe* constructions in Turkic languages are used to express relative characteristics abstracted from the meaning of the attributing nominal component; this tool significantly reduces the need for adjective-noun phrases. In cases of available relative adjectives in Tatar (all of them borrowed, as we mentioned above), we see the following very frequently observed correspondence of grammatical patterns of noun phrases (abbreviations: N – noun, ADJ – adjective, POSS_3 – possessive affix, 3d person):

ADJ + N and *N + N, POSS_3*:

икътисади кризис (ADJ + N), *икътисад кризисы* (N + N, POSS_3)
‘economic crisis’.

Other regular correspondences follow the models *N, NMLZ + N, POSS_3* and *N, PL + N, POSS_3* (abbreviations: NMLZ - nominalization, PL - plural):

акционерлык жәмгыяте (N, NMLZ + N, POSS_3), *акционерлар жәмгыяте* (N, PL + N, POSS_3)
‘joint-stock company’.

Such regular correspondences multiply the number of grammatical variants of multiword terms. In cases where structural (syntactic) variation of synonymous multiword items is complicated by synonymy of their components, we find ramified sets of synonymous items and their derivatives (see Table 6).

Table 6. Synonymous items designating *Constitutional court* and their frequency of use in corpus collections

Compound term	Structure of compound	Frequency in Corpus of Written Tatar	Frequency in Tatar National Corpus
Конституция суды	N + N, POSS_3	905	119
Конституцион суд	ADV + N	169	66
Конституция мәхкәмәсе	N + N, POSS_3	261	126
Конституцион мәхкәмә	ADV + N	48	19

Parallel denominations can be used in a wide range of cases, including official names of state structures and social institutions. So in thesaurus compilation a separate task is to build the most complete list of synonymous denominations related to the same concept to fix them as lexical entries for the thesaurus, because the lists of synonyms in available dictionaries are incomplete (in many cases only one translation variant of a term is represented).

6 Conclusion

Collecting Tatar synonymous items for the Russian-Tatar bilingual thesaurus and analyzing lexical data allows us to conclude that the abundance of absolute synonyms in Tatar is a peculiarity of the Tatar language in its current stage. In the late 1980s a more active social impact on the language situation in Tatarstan began; the Language Law significantly raised the status of the Tatar language, the social movement for purifying Tatar from redundant Russian and Western elements took place, and nowadays the language situation in Tatarstan remains unstable.

The emergence of a great number of exact synonyms in Tatar is caused by a combination of intralinguistic and extralinguistic conditions. The location of Tatar culture at the intersection of Occidental and Oriental civilizations leads to active lexical borrowing both from Arab-Muslim cultural area and European cultural area; borrowing vocabulary from European languages is carried out through mediation of the Russian language, and, certainly, a huge number of words and constructions are taken from Russian. Besides, a significant number of synonyms are built from Turkic and Tatar lexical material. Therefore, words of different origin (Turkic and Tatar, Arabic and Persian, Greek, Latin, English and Russian), coexisting in Tatar, are related to the same referent, which causes a great number of synonyms at the single word level.

Some peculiarities of lexical, derivational and grammatical systems of the Tatar language also lead to the originating of a great number of synonyms. On the level of multiword terms and phrases, lexical synonymy is complicated by the factor of using different grammatical structures, the most characteristic of these being the parallelism of *ADJ + N* and *N + N*, *POSS_3* constructions. As a result, parallel denominations are used in official and media texts for a wide range of phenomena.

Coexistence of different centers of building Tatar terms and the lack of coordination among them leads to instability and imbalance of terminology. We can conclude that a certain degree of Tatar terminology management is thus required, and this may be descriptive (describing how terms are used in documents and media texts) and prescriptive (or even normative) for document compilers (prescribing what terms must be used in standard work and official documentation and how they must be used).

Corpus data analysis proves that synonymy both in socio-political terminology and in general lexicon are characterized by dissimilar features. Absolute synonyms are mainly related to socio-political and scientific terminology (they are characterized by unambiguousness and have a content which in many cases may be well defined) and have direct correspondences in Russian (a Russian term may have a set of Tatar equivalents). Synonymy in general lexicon has a more intricate structure, and the number of absolute synonyms is very limited. A comparative study of synonymy in socio-political domain and in general lexicon is intended as the next step of this research.

References

- Corpus of Written Tatar*. Accessed at: <http://www.corpus.tatar/en> [28/03/2018].
- Cruse, D. Alan. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.
- Desagulier, G. (2014). Visualizing Distances in a Set of Near-Synonyms. In *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*, 43, pp. 145-178.
- DiMarco, C., Hirst, G., & Stede, M. (1993, March). The Semantic and Stylistic Differentiation of Synonyms and Near-Synonyms. In *AAAI Spring Symposium on Building Lexicons for Machine Translation*. Stanford, CA, pp. 114-121.
- Divjak, D. (2010). *Structuring the Lexicon: A Clustered Model for Near-Synonymy*. Walter de Gruyter.
- Divjak, D. (2006). Ways of Intending: Delineating and Structuring Near-Synonyms. In St. Th. Gries, A. Stefanowitsch (eds.). *Corpora in Cognitive Linguistics*. Berlin: Mouton, pp. 19-56.

- Edmonds, P., Hirst, G. (2002). Near-Synonymy and Lexical Choice. In *Computational Linguistics*. 28(2), pp. 105-144.
- Egan, R.F. (1984). Synonym: Analysis and Definition. In Ph. B. Gove, (ed.) *Webster's New Dictionary of Synonyms*. Springfield, Massachusetts: Merriam-Webster, pp. 23a-25a.
- Galieva, A., Nevzorova, O. & Yakubova, D. (2017). Russian-Tatar Socio-Political Thesaurus: Methodology, Challenges, the Status of the Project. In G. Angelova et al. (eds). *International Conference Recent Advances in Natural Language Processing, 2 - 8 September, 2017*. Varn: INCOMA Ltd., pp. 245-252.
- Geeraerts, D. (2010). *Theories of Lexical Semantics*. Oxford: Oxford University Press.
- Inkpen, D., Graeme H. (2006). Building and Using a Lexical Knowledge-Base of Near-Synonym Differences. In *Computational Linguistics*, 32(2), pp. 223-262.
- Lyons, J. (1995). *Linguistic Semantics*. Cambridge: Cambridge University Press.
- Loukachevitch, N., Dobrov, B. (2014). RuThes Linguistic Ontology vs. Russian Wordnets. In *Proceedings of the Seventh Global Wordnet Conference*. Tartu: University of Tartu Press, pp. 154-162.
- Miller, G.A., Charles, W.G. (1991). Contextual Correlates of Semantic Similarity. In *Language and Cognitive Processes*, 6(1), pp. 1-28.
- Miller, G.A. (1995). WordNet: A Lexical Database for English. In *Communications of the ACM*, 38(1)1, pp. 39-41.
- Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D. & Miller, K.J. (1990). Introduction to WordNet: An On-line Lexical Database. In *International Journal of Lexicography*, 3(4), 235-244
- Murphy, M.L. (2003). *Semantic Relations and the Lexicon: Antonymy, Synonymy and Other Paradigms*. Cambridge: Cambridge University Press.
- Nevzorova, O., Mukhamedshin, D. & Bilalov R. (2015). Search Engine for the 'Tugan Tel' Tatar National Corpus: Main Decisions. In *Proceedings of the International Conference "Turkic Languages Processing TurkLang-2015"*. Kazan: Academy of Sciences of Tatarstan Republic Press, p. 236-244.
- Niina N., Nina P. (2010). Synonymy in Specialised Communication – a Terminological Approach. In *Re-thinking synonymy: semantic sameness and similarity in languages and their description*, book of abstracts, pp. 66-67.
- Russian-Tatar Socio-Political Thesaurus*. Accessed at: <http://tattez.antat.hru> [28/03/2018].
- RuThes Linguistic Ontology*. Accessed at: http://www.labinform.ru/pub/ruthes/index_eng.htm [28/03/2018].
- Suleymanov, D., Nevzorova, O., Gatiatullin A., Gilmullin, R. & Khakimov, B. (2013). National Corpus of the Tatar Language "Tugan Tel": Grammatical Annotation and Implementation. In *Procedia - Social and Behavioral Sciences*, 2013, 95, pp. 68-74.
- Taylor, J. R. (2002): *Cognitive Grammar*. Oxford: Oxford University Press.
- Tatar National Corpus. Accessed at: (<http://tugantel.tatar/?lang=en> [28/03/2018]).
- Ullmann, S. (1962). *Semantics, An Introduction to the Science of Meaning*. Oxford: Basil Blackwell.
- Vossen, P. J. T. M. (ed.) (2002). *EuroWordNet: General Document*. Amsterdam: University of Amsterdam.

Acknowledgements

The reported study was funded by Russian Science Foundation, research project No. 16-18-02074.