Semantic Distance Coefficient for Semantic Field of Tatar Culture Specific Concepts "Family and Human Being Environment" and Their Lacunas in English

Gulnara M. Nurtdinova[1]

[1] Kazan (Volga Region) Federal University, Kazan, Russia

Correspondence: Gulnara M. Nurtdinova, Kazan (Volga Region) Federal University, Kremlyovskaya Street 18, Kazan 420008, Russia. E-mail: gnurtdinova@bk.ru

Abstract

The importance of the issue is caused by the demand to translate literature and periodicals from Tatar into English. While translating we can have difficulties with the lexemes that do not have equivalents in English or their meaning differ in some nuances so such lexemes can be bottlenecks of cross-cultural communication. The aim of the article is to research the semantic group of Tatar culture specific concepts of the topic "Family and Human being environment" that do not have equivalents in English. Research methods we used are the method of a dictionary article analysis, comparative and analytical methods. The research results showed that the semantic distance coefficient for semantic group of Tatar culture specific concepts and their equivalents "Family and Human Being Environment" made up 8.57 and it is showing that the sphere of concepts of this field differs much for Tatar and English people. The research results can be used in future research of a Language vitality problem as well as in Theory and Practice of Machine Translation.

Keywords: language worldview, vitality, equivalent, lexeme, realia, lacuna, coefficient of semantic distance

1. Introduction

1.1 The Status of the Problem

In present time, the time of the world globalization, the problem of languages preserving is one of the issues of the day. The languages preserving is closely connected with their vitality. M. Meyerhoff in her book 'Introducing Sociolinguistics' is explaining the reasons why some languages are becoming world ones and why the others are disappearing. In her opinion three pillars represent vitality as being a function of three factors: the status (variety of the speakers of that variety in different contexts), the demographics of the group identified and identifying with that variety, and institutional measures supporting or recognizing a variety. The rise of the language status depends on a number and a status of people who speak it (Meyerhoff, 2011). At present time English has the status of a world language and one of its functions is to bridge the world peoples (Proshina, 2002). The number of people speaking English is the biggest in the world and the authorities in all the countries support its learning. So English can be the bridge of communication and it can promote learning the cultures and literature of other countries' peoples. That is why many pieces of literature and folklore of different peoples are translated in English. The status of Tatar language is high on the land of the Republic of Tatarstan, the half of the population speaks it and the authorities are supporting its learning. Moreover, some foreign countries, such as Australia, Finland, America, and China have Tatar Diasporas where Tatars live and speak Tatar and try to preserve their native language. Unfortunately the status of Tatar in the world is not so high because the number of people speaking it is small and one more reason of its low status can be explained by the fact that very few work pieces of literature and folklore are translated in English from Tatar. So we can say that a small number of people have the information about Tatar people culture and literature and consequently, we can speak about lack of information of Tatar language and culture that is proved by a small number of published Tatar-English and English-Tatar dictionaries and translated works of Tatar writers. To solve the above-mentioned problem we want to compare lexical systems of Tatar and English, particularly, the semantic field "Family and Human being environment". We researched the corps of Tatar fairy tales collections written in 1939-1940 and selected the Tatar lexemes that do not have equivalents in English. We tried to explain the reasons of lexicalization lack for

these lexemes and presented the results.

1.2 Research Materials and Methods

The object of our research is the collection of Tatar fairy tales. Some of them were taken from the collection of Tatar fairy tales of Kaum Nasyry, famous Tatar educator and writer, published in 1900 and the others were taken from the collection of Tatar fairy tales that were written down by well-known Tatar writers Gumer Bashirov and Khamit Yarmyand during the ethnographic expeditions in 1939-1940. Basing on the collections comments we can conclude that the tales were told and then written in so-called "a village koine", the language received by joining Tatar language dialects, it was also the source for "a town koine" development (the stages of Tatar language development are presented in Tatar Grammar, 1992: volume 1). Any fairy tales are written in the style of folklore and the Great Russian writer M. Gorky wrote that the source of a language can be found in folklore. He recommended to collect the folklore, learn and research it. It gives much for writers and poets (Gorky, 1953). The choice of research material is caused by the fact that the language of the tales has many specific Tatar culture concepts which are the subject of our research.

The Russian linguist S. Ter-minasova has presented the links between worldview and culture and language worldviews in her book "Language and Intercultural Communication" (Ter-minasova, 2000). In her opinion any language verbalizes national culture worldview; any language preserves the culture and passes it through generations. She is giving the definition showing that any language is a powerful social instrument that can direct people in ethnos to found the nation through preserving and passing the culture, traditions, social consciousness of a particular language community (Ter-minasova, 2000). She is expressing the idea that a language worldview reflects the universe world image and culture and language world views are closely connected and they are originated from the universe worldview. Consequently while learning the language we learn the culture. Any language is a complex system which is divided into some subsystems. One of language subsystems is a language lexicon. A lexicon is closely connected with external linguistic factors that cause the language change due to any changes in social environment. The researcher A. Lomov considers that modern linguistics is aimed at researching external links of a language with the universe because a language reflects it (Lomov, 1996). Many linguists have conducted research in this field and one of the founders of this branch was Swiss linguist F.Sossyur.

One of the main lexicon units is a word. A word expresses the relation to an object naming some compartment of the universe reflecting it through the concepts of a particular ethnic community. As the concept forming ways are different for different peoples, by some reasons, one and the same universe fragment can be interpreted differently causing semantic discrepancy which are observed in the process of language and culture contacts. Foundation of a new society or invention of something new cause forming new words in the language and at the same time some words transfer to archaisms or get new connotative meaning.

Russian researchers E. Vereshshagin and V. Kostomarov consider that all language units including phonemes and syntax models possess the form and the meaning (Vereshshagin & Kostomarov, 1990). They use the terms 'lexeme' and 'lexical concept' to denote a word form and a word meaning. In their opinion 'lexeme' is a word naming an object, a situation, a descriptor or phenomenon while 'lexical concept' is a set of descriptors with the help of which people can realize if they can use this word to name the object. Thus, lexical concept can be interlanguage one: *book (English), книга (Russian), китап (Tatar)* and it can be translated in other languages because it has direct equivalents in many languages and there is no need for other ethnic group to form a new word. E. Vereshshagin and V. Kostomarov name such interlanguage concepts equivalent words. But lexicon of any language has the words lexical concepts of which do not have equivalents in other languages. Such words some researchers call non-existing or culture-specific words. The number of culture specific words is very small in any language; the results of the research conducted by E. Vereshshagin and V. Kostomarov showed that culture-specific words in Russian made up 6-7% (Vereshshagin & Kostomarov, 1990) but most culture-specific words can be the source for loan words, they can enrich the lexicon and identify the culture a person belonging to.

The words that do not have equivalents in other languages have been researched by many linguists presenting various branches of linguistics: comparative linguistics, translation studies, theory of intercultural communication, ethnopsycology, folkloristics and others. They tried to classify such words and one of the classifications was made by Russian researcher L. Barchudarov (Barchudarov, 1975). He is dividing such words into three groups: 1. Proper and geographic names. 2. Realias, the words that do not exist in the language of people of other community. Usually such words present specific culture concepts of particular ethnic group. 3. Lexemes that L. Barchudarov named occasional lacunas. He gave such a name for the words that do not have

equivalents in other language for some reason: *pogorelets (a person whose house has been burned), kipyatok (boiling water),* no equivalents in English and Tatar *(Barchudarov, 1975).* We are researching the second group, Tatar culture specific words that do not have equivalents in English basing the opinion of E. Vereshshagin and V. Kostomarov (Vereshshagin & Kostomarov, 1990) that culture specific words cannot have the equivalent in language A but they can have it in language B: light blue color (goluboi) has a special world in Russian but does not have the equivalents in Russian and Tatar.

2. Results

2.1 Theory Review

Many linguists have been conducting research in the field of non-existing words and they tried to give their own definitions and terms such as realia, lacuna, lexical gap but the problem of the definition is not solved yet. In opinion of N. Fenenko (she uses the word realia to denote culture specific notion) this fact can be explained that the term *relaia* was taken from other sciences (Fenenko, 2013). In her monography she is giving her own definition where she is defining realia as a source concept and to preserve it as a pair term for lacuna, the notion of comparative linguistics. Realia is denoting the significate of a word relating to one lingual culture when this word does not have nomination in other language and the lexical gap can be named "lacuna" (Fenenko, 2013). She is writing that the analysis of the existing definitions of the term "realia' have shown undifferentiated character of its use.'Realia' is used both to name the culture object of reality and for the language nomination of the same object that causes the misunderstanding of the term interpreting. In this connection she proposes to summarize the opinions of researchers in the field of translation study and comparative linguistics who differentiate the use of the term "realia" in the following way: on the one hand, they denote objects related to history, culture, economy with the word "realia"; on the other hand, they denote the names of such objects by the words-realias, names-realias and terms-realias. Here we have chosen the term 'realia' to name culture specific concepts.

One more problem of the notion "realia" is a criterion basing on which we can refer the word to the group of culture specific words or realia. N. Fenenko is considering that culture specific words form a subsystem of the lexicon of a particular language. She is giving such a group the name "realicon". Basing the method developed by the researchers of Voronezh University, Russia and presented in the collection of articles (N. Feneko, A. Kretov, I. Bulgakova and others, 2013) the criterion of the realicon selection can be absence of the word's equivalent in bilingual dictionary or the equivalent of the word can be presented by a few words in metalanguage. It should be noted that the definition have to be a free word combination because a set-phrase can be interpreted as a language unit and cannot be accepted.

2.2 Tatar Realias and Their Equivalents Presented by Lacunas in English

We have researched Tatar fairy tales to select Tatar culture specific concepts namely, one of their groups, relaias, the words that do not have equivalents in English. Such words verbalize some national culture fragments of the universe. Their lacunas in English present the difference of geographic and historical conditions in which Tatar and English people live as well as the difference in culture worldview. As it was mentioned above the criterion of selection is the absence of equivalent in bilingual dictionary or a definition presented by a free word combination in metalanguage. We researched dictionary articles of printed English-Tatar dictionaries, e-dictionaries and Explanatory dictionaries of English and Tatar. The researchers L. Barchudarov (1975), E.Vereshshagin and V. Kostomarov (1990), S. Vlachov, S. Florin (1980) consider that realias can be mostly nominative language units, nouns, very rarely, adjectives which were formed from nouns denoting specific notions. Thus we are researching the lexemes of Tatar that do not have direct equivalents in English.

We researched forty eight fairy tales and got 579 realias. Basing the method used by N. Fenenko (Fenenko. 2006), at the first stage of the research all the words-realias should be distributed into semantic fields and groups. The difference in lexicalization of objects and the phenomena of the environment is getting obvious while comparing the same semantic groups of two or more languages (Fenenko. 2006). First we selected the realias from the corpus of Tatar fairy tales, excluded the loan words and distributed them into the following different semantic groups (Nurtdinova, 2014):

1) Family and Human being environment.

2) Animals.

3) Household objects

4) Religion.

5) Landscape.

6) Fairy tales characters.

7) Occasional lacunas (see above).

In our article we are analyzing the first group "Family and Human Being Environment".

2.3 Family and Human Being Environment

The results we have got show that Tatar people has more detailed family hierarchy in comparison with English people. In Tatar some word referring to family do not have direct equivalents: *бикәч* young wife, *кәрдәш* relative, as an address, *бертуган* native relative, *сенел* younger sister, *кодагый* mother of a daughter/son-in law toward their parents or any other elder relatives, *бажай* brother of a wife's husband, *эне* younger brother toward an elder brother/sister, *каене* younger brother of a wife, *бабалар* wife's parents*, эти, әни* father and mother-in-law, *җиңгә* elder brother or any elder relative's wife. Tatar people also differentiate between elder and younger sisters and brothers while addressing them: *апа, түтәй* elder sister, *сенел* younger sister, *абы* elder brother, *эне* younger brother. Tatars address any older person naming them *абы* and *апа (*the words mean elder brother or elder sister). In English given concepts are not lexicalized by one word, their definitions are presented by few words proving that absence of a lexeme while the phenomenon is existing show that the phenomenon is not so important for people of a particular culture. The list of close relatives is much narrower in English than in Tatar and the age of relatives is also important for Tatars that is not presented in English.

2.4 Human Being Environment

The analysis of the group showed the following lacunas in English: *түбә* place, where a few relatives live nearby, *абыстай* mullah's wife, *бай* landlord, rich cattleman hiring farmhands, *буре куучы* wolf hunter as a profession, *кәрлә* bearded short man or bearded pygmy, *чулак* one-armed or one-handed person, *кәнизәк* maid of a rich lady working for the lady only, *кәләшлек* the status of engaged bride, *хатынлык* the status of a wife, *аксак* lame person and others. The cited Tatar lexemes and their lacunas in English present the phenomena not so important for English people in comparison with Tatars, all they need the explanation. In particular, they show that English people do not pay much attention to appearance (*кәрлә* bearded short man or bearded pygmy*, чулак* one-armed or one-handed person*, аксак* lame person). The status of a married woman is presented by the lexemes wife/spouse in English when Tatars pay attention to the status and the age of a wife (*кәләшлек* the status of engaged bride, *хатынлык* the status of a wife, *абыстай* mullah's wife). The lexeme *түбә* place, where a few relatives live nearby one more time presents closer relationships of relatives of Tatar people. Often the residents of Europe complain of too many relatives who come together to live in their region. One of the reasons of misunderstanding is that culture of some peoples supposes closer relationships between relatives.

2.5 Coefficient of Semantic Distance for the Semantic Group "Family and Human Being Environment"

It is well known that it is impossible to learn the language not comparing it with the other one. So the next step of our research is to calculate the coefficient of semantic distance for the semantic group "Family and Human Being Environment" for Tatar lexemes that do not have equivalents in English. The higher is the figure the bigger is the distance between the concepts. The method was proposed by Russian researcher V. Nalimov (Nalimov, 1979) to describe the meaning of explicated word and later it was used by V. Titov (Titov, 2002) to calculate the coefficient of semantic distance for a pair of any languages. According to this method the criterion for a lexeme selection can be the following condition: if a word of any language has an equivalent in an explanatory article of the dictionary it means that both languages segment the semantic field in similar way and form the similar concepts. But in the case when to explain the word meaning we need a few words we can say that the pair language does not have analogous concept. V. Titov is concluding that the more words we need to explain the word meaning the less common concepts the pair languages have (Titov, 2002).

To calculate the coefficient of semantic distance it is necessary to get the data about the number of the definitions from a bilingual dictionary and the number of words in the definitions. Then we should get the total number of the words of all the definitions. To do it we have to multiply the number of the words of each definition by number of the definitions and total number we got will be the quantity D – the sum of all the words of the definitions. V. Titov is presenting the following formula of a coefficient of semantic distance calculation for a particular language: a coefficient of semantic distance = D : Q, where Q is the number of all the definitions in the dictionary).

The researcher N. Fenenko have used this method to calculate the coefficient of semantic distance for French and Russian concepts and then she did the same for French realias selecting them from bilingual e-dictionary. Basing the data she got the coefficient of semantic distance for French and Russian is 1.73 when the coefficient for

French realias is 8.1475. The stated figures show that culture specific concepts correlate with nearly 100% of lacunas and demand special research to bridge alien with own culture. She is making the conclusion that to decode culture specific concepts we always need more than one word if not we can speak about lexemes that have straightforward equivalents and exist in many languages (Fenenko, 2006).

We selected forty three realias of the semantic group "Family and Human being environment" (selected from Tatar fairy tales) to explain them it has been used from two up to nineteen words. Basing the above mentioned opinion we excluded Tatar lexemes that have straightforward equivalents in English.

Here we are presenting the draft fragment of Tatar English dictionary for Tatar culture specific notes. We are using the method proposed by V. Titov for calculating the coefficient of semantic distance for Romanic languages. He is giving the following example: to decode the meaning of Romanian word *portbagaj* in Russian we need 11 Russian words: *n. сетка для вещей (в вагоне); багажник м. (мотоцикла и т.п.), 11 (Titov, 2002)*. We applied the method for Tatar concepts.

Әби *n. grandmother, mother-in-law, wife's mother, grandma, old lady, old woman 9*

Бабай *n. grandfather, father-in-law, wife's father, grandpa, old man, old chap 10*

Әти *n. father, father-in-law, husband's father, papa 5*

Әни *n. mother, mother-in-law, husband's mother, mama 5*

Әдәм бала *mankind, human family, human race, Adam's son 7*

Агай *(col.) elder brother, brother 3*

Түбә n. *top, hill, ceiling, roof, sky, dome of heaven, relatives living in neighborhood 11*

Бикәч n. *young wife 3*

Кәрдәш *n. address to anyone, relative, member of a family 9*

Абыстай *n. mistress, priest's wife, wife of any rich man* 9

Келәш (dial.) *n. young wife, giggling girl 4*

Бай n. *landlord, any rich man, address to a rich man 9*

Патша *(фадиша, dial.) n. padishah, title formerly used for various rulers, including the shah of Iran and the sultan of Turkey, king 19*

Хан *n. khan, title given to rulers and officials in central Asia, Afghanistan, and certain other Muslim countries 17*

буре куучы *wolves hunter 2*

Ат *n. horse, family (arch.), stock 5*

Мулла *n. mullah, moollah, Muslim teacher of law and religion, Muslim learned in Islamic theology and sacred law17*

Кәрлә *n. short bearded man, gnome, dwarf, bearded pygmy 8*

Бертуган *n. full sister/brother in immediate/nuclear family 8*

Сеңел *n. full younger sister, address to any girl or woman who is younger 12*

Таз *n. scabby man, mangy man, (in the fairy tales) wheeler and dealer 12*

Энекәш *n. address to any man who is younger, younger brother in immediate family 13*

Җиңгә *n. wife of elder brother or any elder relative 9*

Кодагый *n. mother of son-in-law; mother of daughter-in-law for wife's or husband's parents, elder sisters or female relatives of fiancée 19*

Тута *n. elder sister, address to the elder sister in immediate family, noble lady, princess 14*

Җамавыз *n. widow, any old woman living alone 7*

Эне *n. younger brother in immediate family, full brother 8*

Чулак *n. armless or one–armed man, left-hander, left-handed person 8*

Бажай *n. brother-in-law, husband of wife's brother 6*

Кәнизәк *(hist.) n. servant for rich lady, maid of honor 9*

Каене *n. wife's younger brother, husband's younger brother, brother –in-law 8*

Кортка *(col.) n. widow, wife, old woman 6*

Кәләшлек *n. to be in the status of fiancée, to be engaged 11*

Җәмәгать *n. society, wife (arch.), community 5*

Хатынлык *n. to be in the status of wife, weak-willed, lack backbone 13*

Юлдаш *n. fellow traveler 3*

Ахмак *(col.). n. fool, silly 4*

Бабалар *n. father and mother – in law for wife's parents 8*

Багаучы *(hist.) low rank official 4*

Аксак *n. lame man, limping 4*

Вәзир *n.vizier, high official in some Muslim countries 8*

Әмир *n. emir (amir), title of various Muslim (mainly Arab) rulers 9*

Атаман *n. cossack leader, leader in plays, gang's leader 8*

Total – 43

Table 1. Family and human being environment

Number of words	Number of the definitions	total
2	1	2
3	3	9
4	4	16
5	4	20
6	2	12
7	2	14
8	7	56
9	7	63
10	1	10
11	2	22
12	2	24
13	2	26
14	1	14
15	-	-
16	-	-
17	2	34
18	-	-
19	2	38
	Q - 42	D - 360

coefficient = 360/42=8.57

Thus the coefficient of semantic distance for realias of semantic group "Family and Human being environment" is 8.57. The data is showing that Tatar culture specific concepts differ much from English in the semantic area of family relations and translators or lexicographers should take it into consideration to avoid misunderstanding.

3. Conclusion

Lexicon of any language has culture specific concepts that do not have straightforward equivalents in other

languages. Such concepts can be considered as bottlenecks in the process of intercultural communication and they mark alien and own culture. Many linguists have been conducting research of these concepts and define them in their own way but still there is no a united system for research and that is why the researchers face the problem of the variety of the terminology and consequently the methods. In this article we are basing the terminology and the method proposed by Russian researchers L. Barchudarov (1975), V. Titov (2002) and N. Fenenko (2006-2013). We accepted the system of the words that do not have straightforward equivalents developed by L. Barchudarov. He has divided such words into three groups. 1. Proper and geographic names. 2. Realias or the words naming the culture specific concepts that do not exist in the experience of people speaking other language. 3. Occasional lacunas (Barchudarov, 1975). We researched the second group of words, culture specific words. Our research was based on the hypothesis developed by the representatives of Voronezh university of Russia. N. Fenenko, one of the researchers of the university uses the terms 'realia' for such concepts and 'lacuna' to name a lexical gap of the concept in other language. Basing the method and the criteria she used for selecting relias and their lacunas in English we researched Tatar culture specific concepts taken from Tatar fairy tales. We selected all the lexemes that do not have straightforward equivalents in English and distributed them into different semantic groups. Then we calculated the coefficient of semantic distance for semantic field "Family and Human Being Environment". The results showed that Tatars and English differ much in the area of family hierarchy that can cause misunderstanding in the process of intercultural communication while the world community is learning Tatar literature and culture. The coefficient for this group made up 8.57.

4. Recommendations

The method of the research that is described in the article can be used to compare semantic groups and fields aimed at revealing the difference in the process of lexicalization of world culture and language views via any language. The research can be the basis for Tatar English Explanatory and bilingual dictionaries of Tatar culture specific concepts and it can contribute to the development of machine translation for Tatar and English languages. World community will have the opportunity to learn Tatar culture and promote the vitality of Tatar.

Acknowledgments

References

ABBYY Lingvo x3 dictionary.

Amirova, T. A. (2006). *Istoriya yazykoznaniya* (3rd ed.). Moscow: Academia.

Barchudarov, L. C. (1975). *Yazyk & Perevod.* Moscow: Mezdunarodnye Otnosheniya.

Fenenko, N. A. (2006). Franzhuzskiye realii v kontekste teorii yazyka ((Doctoral dissertation). University of Voronezh, Russia.

Fenenko, N. A. (2013). *Franzhuzskiye i russkiye relii v aspecte teorii mezhyazykovoi renominatsii.* Voronezh: Izdatelstvo Voronezhskogo Universiteta, Russia

Gorky, A. M. (1953). Sobraniye sochineny. Volume 24. Moscow.

Inglizcha-tatarcha suzlek. (2007). Kazan: Magarif, Russia.

Lomov, A. M. (1996). Linguistika & analiticheskaya philosophiya. Vestnik VGU. Ser. 1. Gumanitarnye Nauki (76-85).

Longman Dictionary of Contemporary English – URL: http://www.ldoceonline.com

Meyerhof, M. (2011). *Introducing Sociolinguistics* (2nd ed.). Routlegde, London and New York.

Nalimov, V. V. (1979). *Veroyatnostnaya model yazyka: O sootnoshenii yestestvennych i iscusstvennych yazykov.* Moscow: Nauka.

Nutrtdinova, G. M. (2014). Vestnik CHGU. Philologiya. *Iskusstvovedeniye, 91*(102-108). 1994-2796.

Proshina, Z. G. (2002). Anglisky kak posrednik v communokatsi narodov Vostochnoy Asii I Rossii: problema oposredovannogo perevoda (Doctoral dissertation). University of Vladivostok, Russia.

Tatar telenen alnatmaly suzlege. (2005). Kazan: Matbugat jorty nashriyate, Russia.

Tatarskaya grammatika. (1992). Kazan: Tatarskoye knizhnoye izdatelstvo. Tom 1. Russia.

Tatarskiye narodnye skazki. (1958). Vol.I. Kazan: Tatarskoye knizhnoye izdatelstvo, Russia.

Ter-minasova, S. G. (2000). *Yazak i mezhkulturnaya kommunikatsiy.* Moscow: Slovo.

Titov, V. T. (2002). *Obshshaya kvantativnaya lexicologiya romanskich yazykov.* Voronezh: Izdatelstvo Voronezhskogo Universiteta, Russia.

Vereshshagin, E. M., & Kostomarov, V. G. (1990). Yazyk i Kultura. Moscow.

Vlachov, S., & Florin, S. (1986). *Neperevodimoye v perevode* (2nd ed.). Moscow: Vysshaya Shkola.

Zakiev, M. Z. (1993). *Tatarskaya grammtika.* Kazan: Tatarskoye knizhnoye izdatelstvo, Russia.