Аксиологический и культурный код
семантики глагола в контексте машинного перевода
(на материале русского, татарского и английского языков)

Айдарова А. М., Билялова А. А., Зеленина Т. И.

Аннотация. Целью данного исследования является выявление аксиологического и культурного кода семантики глагола в контексте машинного перевода на материале русского, татарского и английского языков. Аксиологические и культурные коды оказывают значительное влияние на смыслы, передаваемые в разных языках. Эти наборы кодов являются важными аспектами языковой семантики, особенно в сфере машинного перевода. Научная новизна исследования заключается в изучении глаголов русского, татарского и английского языков с последующей реконструкцией и описанием поведенческого кода носителей данных языков, который представлен существующими нормами, запретами и императивами в названных языковых сообществах. В результате работы раскрываются нормы и ценности, присущие семантике поведенческих глаголов. Полученные данные исследования показали, что люди из разных языковых сообществ разделяют схожее понимание концепта человеческого поведения. В то же время результаты указывают на уникальные аспекты отражения культурных поведенческих кодов в исследуемых языках. Эти коды облегчают распознавание разнообразных коннотаций или нюансов в разных культурах, тем самым способствуя культурно релевантным переводам. Лингвистические и культурные элементы, выявленные посредством сравнительного анализа, имеют решающее значение для точного представления слова в словарях, тезаурусах, базах данных, текстовых корпусах, особенно при машинном переводе, в ходе которого незнание специфического аксиологического кода, отражающего систему ценностей конкретной культуры, может привести к когнитивному разрыву, интерпретационной ошибке и даже культурному конфликту. Внимание к аксиологическим аспектам семантики глаголов является существенным для разработки более продвинутых и чувствительных к культурным различиям систем машинного перевода.

Axiological and cultural codes
of verbal semantics in the context of machine translation
(based on the verbs of the Russian, Tatar and English languages)

A. M. Aydarova, A. A. Bilyalova, T. I. Zelenina

Abstract. The aim of the research is to identify, based on verbs of the Russian, Tatar and English languages, axiological and cultural codes in the context of machine translation. Axiological and cultural codes hold considerable sway over the meanings conveyed in various languages. Both sets of codes are vital facets of language semantics, particularly within the realm of machine translation. The scientific originality of the research lies in the study of the verbs of the Russian, Tatar and English languages, followed by the reconstruction and description of the behavioral code of the speakers of these languages, which is represented by the norms, prohibitions and imperatives existing in these linguistic communities. As a result, the article delves into the norms and values inherent in the semantics of behavioral verbs. The study shows that people from different language communities not only perceive the concept of human behavior in the same way, but also convey the cultural features of that behavior differently in their language. This means that although the basic ideas about human behavior may be similar, each culture has its own unique codes of expressing these ideas through language. Linguistic and cultural elements identified through comparative analysis are crucial for accurate word representation in dictionaries, thesauri, databases, and text corpora, especially in machine translation, where ignorance of a specific axiological code reflecting the value system of a particular culture can lead to cognitive discontinuity, interpretive error, and even cultural conflict. Attention to axiological aspects of verb semantics is essential for the development of more advanced and culturally sensitive machine translation systems.

Introduction

The relevance of this study consists in the increased interest in cross-cultural communication and translation techniques for distant languages. In this regard, studies in which the language is considered as a carrier of the cultural and axiological code of a particular ethnic group remain topical and urgent in modern linguistics. In the context of machine translation, these codes play a significant role in ensuring accurate and culturally appropriate translations.

The tasks of the research are as follows:

1) to compile a list of behavioral verbs in the Russian, Tatar and English languages that bear axiological and cultural codes in their semantics;

2) to present in table form the results of the comparative semantic analysis of the verbs of behavior in the Russian, Tatar and English languages;

3) to find out the common and unique aspects in each language portrayal of cultural behavior codes;

4) to reveal the role of axiological and cultural codes of verbal semantics in the context of machine translation.

In the course of the work, the following research methods were used: the inductive and deductive research methods, heuristic methods related to the collection and documentation of the material under study; the descriptive method in the analysis of the contextual use of the studied units in lexicographic sources and national corpora; the comparative method in the research of evaluative, cultural and axiological components contained in the semantics of lexical units; as well as the statistical method for calculation purposes.

The material of the study was retrieved from the following dictionaries:

- Cambridge Dictionary. https://dictionary.cambridge.org/dictionary/english/;
- Lexico-Semantic Groups of Russian Verbs: Лексико-семантические группы русских глаголов: учеб. словарь-справочник / авт.-сост. Э. В. Кузнецова. Свердловск: Изд-во Урал. ун-та, 1988;
- Oxford Learner's Dictionaries. https://www.oxfordlearnersdictionaries.com/;
- Russian Semantic Dictionary: Русский семантический словарь. Толковый словарь, систематизированный по классам слов и значений / под общ. ред. Н. Ю. Шведовой. 1998. https://www.slovari.ru/default.aspx?p=235;
- Tatar telenen anlatmaly suzlege. 2015-2021. https://suzlek.antat.ru/;
- Thesaurus.com. https://www.thesaurus.com/.

Materials of national corpora were used to clarify the semantics of verbs:

- British National Corpus. http://www.natcorp.ox.ac.uk/;
- Corpus of Contemporary American English. https://www.english-corpora.org/coca/;
- Russian National Corpus: Национальный корпус русского языка. https://ruscorpora.ru/;
- Tatar National Corpus "Tugan Tel". https://tugantel.tatar/;
- Written Corpus of the Tatar Language. https://www.corpus.tatar/.

The theoretical background of this study is reflected in the thematic (ontological) approach to language material. It is well-known that the guiding principle of foreign language linguistics continues to be the generativist approach, which has become the basis of many modern semantic classifications of English vocabulary, in particular, B. Levin (1993), the semantic verb database VerbNet, Kipper et al. (Kipper, Korhonen, Ryant et al., 2006). Representation of verbs at Princeton University Project in WordNet and other similar projects involves highlighting synsets (synonymous series) of verbs and establishing relationships between synsets, while the selection of large semantic classes, united thematically, has not been considered as a task.

The study is based upon the works of domestic and foreign linguists in the field of lexicology and lexical semantics (Апресян, 1995; Levin, 1993); semantics of behavior (Плотникова, 2009); semantic equivalence (Которова, 1997); the axiological and cultural code of semantics (Kyshtymova, Anikeeva, Mochalova, 2020; Rubenok, 2021); machine translation (Almahasees, 2021; Bowker, Buitrago Ciro, 2019; Specia, Scarton, Paetzold, 2022; Varghese, Punithavalli, 2019).

The practical value of the work lies on the fact that description of the predicates of languages with different structures in the context of ethnocultural and social problems makes it possible to fill in some gaps in the field of comparative linguistics. The results of the investigation expand the theoretical base necessary for the development of linguodidactic foundations for teaching these languages as foreign languages, and also for accurate representation of verbs in dictionaries, thesauri, databases, textual corpora and machine translation. In addition, the results of the study can significantly contribute to the creation of teaching materials for teaching disciplines such as linguocultural studies, terminology, and comparative linguistics. Understanding the interaction of linguistic predicates with cultural and social contexts allows teachers to develop more effective curricula that emphasize the nuances of language use across cultures.

Discussion and results

Traditional ideas about normal and abnormal behavior in different cultures and sociocultural groups can differ significantly. The article examines a group of verbs that describe human behavior in three different languages: Russian, English, and Tatar. Generally, these behavioral verbs carry a negative connotation. This negative evaluation can be either explicit or implicit. Actions that deviate from normal behavior are often viewed as wrong and blameworthy. Verbs in these languages tend to condemn qualities such as dishonesty, insincerity, irresponsibility,

and unnatural behavior. Violations of behavioral norms provoke strong emotional, verbal, and cognitive responses across different cultures, as reflected in the meanings of these verbs (Aydarova, Aminova, 2016; 2019).

Machine translation has become widely used throughout the world as a medium of communication between those who live in different countries and speak different languages. Machine translation systems have become increasingly easy to access and use (Bowker, Buitrago Ciro, 2019; Specia, Scarton, Paetzold, 2022). Cross cultural communication and translation between distant languages constitutes some challenges, for machines in particular (Almahasees, 2021; Varghese, Punithavalli, 2019). The semantics of separate words in any language bear axiological and cultural codes, which refer to the values, beliefs, attitudes or the customs, traditions, practices of a particular culture. Therefore, in the context of machine translation, axiological and cultural codes play a significant role in ensuring accurate and culturally appropriate translations, as well as in the process of designing and developing effective machine translation systems.

The study of the types of behavior of an individual has long attracted the increased attention of scientists of various humanitarian areas, including philology, having received the most complete presentation in the works by L. Bloomfield, who created a behaviorist theory in the middle of the 20th century, representing language as a form of human reaction to the influence of the external environment (Sperber, 1945). However, this approach cannot be considered linguistic in its purest form, since it presents to a greater extent the starting points which are characteristic of the psychology of human behavior. An integral linguistic study of lexical units representing human behavior seems to be extremely advantageous against this background.

Behavior verbs form a class of social verbs that attract interest for various reasons, including their linguistic and cultural characteristics, as well as their grammatical and semantic properties. Most of these verbs carry a negative connotation, indicating a breach of societal norms. In the study of different types of moral norm violations in the Russian language, A. M. Plotnikova identifies several subgroups within the lexical-semantic category of behavioral verbs (Плотникова, 2009).

1. Violation of the principle of sincere behavior: *Rus. лицемерить*, *Tat. ikejözlelänü*, *Eng. pretend.*
2. Violation of the principle of natural behavior: *Rus. важничать*, *Tat. kylanu*, *Eng. show off.*
3. Violation of the principle of moderately modest behavior: *Rus. хвастаться*, *Tat. maktanu*, *Eng. boast.*
4. Violation of the principle of serious behavior: *Rus. глупить*, *Tat. akylsyzlanu*, *Eng. monkey around/about.*
5. Violation of the principle of emotional self-control: *Rus. неистовствовать*, *Tat. kotyru*, *Eng. bridle.*
6. Violation of the principle of ethical self-control: *Rus. кутить*, *Tat. tipterü*, *Eng. carouse.*
7. Violation of the principle of volitional self-control: *Rus. артачиться*, *Tat. üzbeldeklelänü*, *Eng. flout.*

Thus, similar behavior patterns are evident in Russian, Tatar, and English, which makes sense because defining negative behavior helps establish what is considered normal. These shared patterns involve linguistic, cultural, and value-related aspects of meaning. Comparative analysis of these components is important for dictionary entries, translation, and language teaching.

In comparative language studies, the concept of interlingual equivalence is crucial. Interlingual equivalents are units from different languages that either fully or partially share the same meaning and can substitute for one another in specific contexts. Lexical-semantic units of each language reflect distinct sets of concepts, which form the unique national identity of that language semantic system.

Let us consider equivalence relationships on the basis of the verbs of unnatural behavior. The group of verbs of unnatural behavior in Russian is represented by 61 units, in Tatar by 48 units, in English by 27 units. In all three languages, this group includes two subgroups: verbs that describe unusual or affected behavior, and verbs that describe self-importance, showiness, or self-exposure. The results of the comparative semantic analysis are presented in Table 1.

Table 1. *Results of the comparative semantic analysis of verbs of behavior in the Russian, Tatar and English languages*

Semes	Russian	Tatar	English
Verbs describing unnatural behavior (violation of the principle of natural behavior)			
I. Group of verbs denoting original, strange, mannered behavior			
'to behave peculiarly, unusually, strangely'	+	+	+
'to behave coyly by making unnecessary gestures'	+	+	+
II. Group of verbs denoting self-importance, panache, exposing oneself			
'to try to show yourself from the favorable side in order to attract attention'	+	+	+
'to behave immodestly, emphasizing in speech and behavior one's merits, dignity'	+	+	+
'to flaunt something as one's advantage, showing arrogance and self-satisfaction'	+	–	+
'to show superiority over someone with your behavior and appearance'	+	+	–
'to consider oneself better than others, expressing one's disdain for someone'	+	–	–
'to boast of one's courage, indifference to danger'	+	+	–
'to boast arrogantly, showing excessive pride'	+	–	+

The English and Tatar languages have a limited number of verb units for describing unusual, unique, or affected behavior. In Tatar, there are only two verbs, *kylanu* and *artistlanu*, which express this type of behavior, and they do not encompass the full range of nuances that Russian verbs do. In English, such behavior is typically expressed by the verb *philosophize*, which corresponds to the Russian verbs *философствовать* and *мудрствовать*, focusing more on a person's verbal behavior.

The verbs of the Russian language can describe in more detail self-importance, panache, flaunting oneself. The Tatar language does not differentiate between signs of behavior verbalized in Russian verbs with the semes 'to show off something as one's advantage, showing arrogance and self-satisfaction', 'to consider oneself better than others, arrogantly expressing one's disdain for someone', 'to boast arrogantly showing excessive pride'. In English, Russian verbs with the semes 'to show superiority over someone with all their behavior and appearance', 'to consider oneself better than others, arrogantly expressing one's disdain for someone' do not find monolexemic equivalents. The verbs of the Russian language are marked with a bright negative evaluative component (indicated by the label 'disapproving' (Russian Semantic Dictionary)) and belong to colloquial, jargon and vernacular vocabulary. The Russian verbs *храбриться*, *геройствовать*, etc. with the seme 'to boast of one's courage, indifference to danger' (Russian Semantic Dictionary) in English correspond to the verbs *bear up*, *brave out*, *psych out*, *strut around*, *swank*, which do not have a negative connotation. Only the verb *swagger about* is marked with the marker 'showing disapproval' / 'usually disapproving' (Oxford Learner's Dictionaries). Among the synonyms of the Tatar language *batyrlanu*, *ätächlänü*, *gairätlänü*, *gairät itü*, *yöräklänü*, *kydrachlanu*, *kyulanu*, the third and last verbs do not contain a clear negative evaluative component (Tatar telenen anlatmaly suzlege). The English verb *flaunt* does not contain a negative evaluative connotation (Oxford Learner's Dictionaries).

In each language, there are different ways of designating one and the same concept. Therefore, when establishing interlingual equivalents and translating Russian verbs of behavior into Tatar and English, one has to use multilevel means, since these languages do not have the corresponding single-word lexical correlates.

For example, in the Tatar language, certain types of unnatural behavior do not have monolexemic verbal correlates. Russian verbs *щеголять*, *козырять*, *фасонить*, *форсить*, *франтить*, etc. are translated into the Tatar language by the analytical verbs *kupshylanyp / koäazlänep yörү* and the phrase *üzen kürsätergä matashu*. The closest equivalents of the Russian verbs *высокомерничать*, *возноситься*, *воображать*, *зазнаваться*, *задаваться*, etc. are the Tatar phrase *üz-üzenne ere totu*, the phraseological units *boryn kütärü / choyu*, *borynny yugary totu*, *ällä kem bulyp kitü*, *büksə (korsak, tüsh) kierü*. The most accurate equivalents of the verbs *гордиться*, *кичиться*, etc. are the phrases *shapyrynyp söyläü*, *gorurlanyp söyläü*. Also, in the Tatar language, there is only one verb that implements the seme 'to behave in a peculiar, unusual, strange way'. Such a meaning in the speech of a native Tatar speaker can also be conveyed by means of different levels. The equivalents of the Russian verbs *мудрить*, *философствовать* are the phraseological units *akyl / safsata / fälsäfä / hikmät satu*. The Tatar correlates of the Russian verbs *оригинальничать*, *отличаться* are the phraseological units *hikmät chygaru / yasau*, as well as the word combinations *aerylyp toru*, *äkämätlär koru*, *yülär syman / əkəmət kylanu*, *үzenchä kylanu / eshläü*.

Certain varieties of unnatural behavior are not nominated in English with the help of one-word verb correlates. The verbs *важничать*, *пыжиться*, *спесивиться*, *чваниться* correspond to the English phraseological units *act high and mighty*, *put on airs*, *assume airs*, *give oneself airs*. The verbs of the Russian language *высокомерничать*, *возноситься*, *воображать*, *зазнаваться*, *задаваться* correspond to the phraseological units *swell with importance*, *get above oneself*.

The given examples indicate that in Russian the semantics of behavior is realized mainly by verbs, while in Tatar and English – by verbal phraseological units, phrasal verbs and verbal word combinations.

Conclusion

The conducted study allows us to outline the following research findings.

The list of behavioral verbs in the Russian, Tatar and English languages with axiological and cultural codes was compiled.

The results of the comparative semantic analysis of verbs of behavior in three languages were presented in table format, which allowed us to reveal interlingual equivalents.

The study findings suggest that people from different language communities share a similar understanding of "human behavior", particularly behavior that goes against moral, ethical, and etiquette standards. The dominance of negatively evaluated behavior verbs supports the idea that there is an imbalance in the language's lexical composition. The component analysis highlights unique aspects in each language's portrayal of cultural behavior codes specific to different ethnic groups. Verbs describing behavior often denote breaches of ethics, morality, etiquette, and law, but by highlighting these violations, they also implicitly define normative human behavior. These verbs suggest that proper conduct involves being natural, sincere, honest, moral, modest, law-abiding, emotionally controlled, cooperative, responsible, and efficient with time. Overall, they imply that one should not violate principles of sincere, natural, modest, serious, emotionally controlled, ethical, and self-regulated behavior.

The linguocultural and value-based components of meaning discovered through comparative analysis are crucial for accurately representing words in dictionaries and translations. This is particularly important for machine translation, a key area in natural language processing. The linguocultural and axiological components of meaning identified in comparative analysis are particularly relevant for machine translation, as it often relies on statistical models and algorithms that may not fully reflect the nuances of language use. This is especially true when it comes to axiological and cultural codes, which are often difficult to quantify and measure. To overcome these difficulties, machine translation systems need to be designed to take into account the axiological and cultural codes of different languages. This could include databases of cultural knowledge or working with human translators who can provide guidance on cultural nuances and appropriate language use. Note also that the semantic analysis of verbs and their processing can serve as a basis for creating and extending databases and thesauri, such as WordNet and VerbNet, where words are grouped into sets of cognitive synonyms (synsets), each expressing a different concept.

As further research perspectives we can identify the study of axiological and cultural codes of other lexico-semantic groups of verbs as well as other parts of speech.

Источники | References

1. Апресян Ю. Д. Избранные труды: в 2 т. Изд-е 2-е, испр. и доп. М.: Школа «Языки русской культуры»; Издательская фирма «Восточная литература» РАН, 1995. Т. I. Лексическая семантика.

2. Которова Е. Г. Проблема межъязыковой эквивалентности в лексической семантике (на основе анализа русского и немецкого языков): дисс. … д. филол. н. М., 1997.

3. Плотникова А. М. Когнитивное моделирование семантики глагола (на материале глаголов социальных действий и отношений): дисс. … д. филол. н. Екатеринбург, 2009.

4. Almahasees Z. Analysing English-Arabic Machine Translation: Google Translate, Microsoft Translator and Sakhr. Abingdon: Taylor & Francis, 2021.

5. Aydarova A. M., Aminova A. A. Representation of Behavioural World Image in Verbal Lexis // Opcion. 2019. Vol. 35. Special Iss. 22.

6. Aydarova A., Aminova A. The Negative Evaluative Component within the Semantic Structure of Behavior Verbs in Russian, English and Tatar Languages // Journal of Language and Literature. 2016. Vol. 7. Iss. 3.

7. Bowker L., Buitrago Ciro J. Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community. Bingley: Emerald Publishing Limited, 2019.

8. Kipper K., Korhonen A., Ryant N., Palmer M. Extending VerbNet with Novel Verb Classes // Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06). Genoa: European Language Resources Association (ELRA), 2006.

9. Kyshtymova I., Anikeeva T., Mochalova Y. Perception of Axiological and Semantic Code of Characters in Animated Discourse // Media Education. 2020. Vol. 60. Iss. 3.

10. Levin B. English Verb Classes and Alternations. A Preliminary Investigation. Chicago – L.: The University of Chicago Press, 1993.

11. Rubenok M. S. Cultural Code from the Perspective of a Conceptual Metaphor // International Scientific and Practical Conference of Young Researcher "Vector 3.0". Proceedings of International Scientific and Practical Conference. Vitebsk, 2021.

12. Specia L., Scarton C., Paetzold G. H. Quality Estimation for Machine Translation. Cham: Springer International Publishing, 2022.

13. Sperber H. Behaviorism in Linguistics // Monatshefte Für Deutschen Unterricht. 1945. Hf. 37.

14. Varghese N., Punithavalli M. Lexical and Semantic Analysis of Sacred Texts Using Machine Learning and Natural Language Processing // International Journal of Scientific & Technology Research. 2019. Vol. 8. Iss. 12.

Информация об авторах | Author information

RU   **Айдарова Алсу Мирзаяновна**[1], к. филол. н., доц.
Билялова Альбина Анваровна[2], д. филол. н., проф.
Зеленина Тамара Ивановна[3], д. филол. н., проф.
[1, 2] Казанский (Приволжский) федеральный университет
[3] Удмуртский государственный университет, г. Ижевск

EN   **Alsou Mirzayanovna Aydarova**[1], PhD
Albina Anvarovna Bilyalova[2], Dr
Tamara Ivanovna Zelenina[3], Dr
[1, 2] Kazan (Volga Region) Federal University
[3] Udmurt State University, Izhevsk

[1] aidalmir@yandex.ru, [2] abill71@mail.ru, [3] zeleninatamara2@yandex.ru

Информация о статье | About this article