ORIGINAL ARTICLE

INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE

ISSN: 2692-515-x, Impact Factor: 9.23
American Academic publishers, volume 6, issue 01,2026

Journal: https://www.academicpublishers.org/journals/index.php/ijai

LINGUISTIC AND LINGUODIDACTIC FOUNDATIONS OF INTERPRETING ETHNOCULTURAL UNITS IN AI-BASED DICTIONARIES

Yusupova Mushtariy Baxtiyor kizi
Karshi State University
Doctorate (PhD) student

ANNOTATION.

This article analyzes the linguistic and linguodidactic foundations of interpreting ethnocultural units in AI-based dictionaries. It identifies the main problems in the representation of culture-specific lexical units in digital lexicographic systems and substantiates the necessity of a multidimensional, culturally oriented model of explanation. The study shows that simplified and universalized definitions do not ensure an adequate understanding of ethnocultural semantics and proposes a set of principles for improving lexicographic description, including cultural markedness preservation, multilevel interpretation and adaptive presentation for different user groups. It is argued that the integration of linguistic, cultural and pedagogical components in AI-based dictionaries significantly increases their educational and intercultural value.

KEY WORDS

AI-based dictionary, ethnocultural units, linguoculturology, lexicography, linguodidactics, cultural semantics, digital lexicography, intercultural competence.

ANNOTATSIYA

Ushbu maqolada AI-lug‘atlarda etnomadaniy birliklarni talqin qilishning lingvistik va lingvodidaktik asoslari tahlil qilinadi. Raqamli leksikografik tizimlarda madaniy xoslangan leksik birliklarning ifodalanishidagi asosiy muammolar aniqlanib, ularni izohlashda ko‘p qatlamli, madaniyatga yo‘naltirilgan model zarurligi asoslab beriladi. Tadqiqotda soddalashtirilgan va universallashtirilgan ta’riflar etnomadaniy semantikani to‘liq ochib bera olmasligi ko‘rsatilib, madaniy belgilanganlikni saqlash, ko‘p bosqichli talqin va foydalanuvchiga moslashuvchan taqdimot tamoyillari taklif etiladi. AI-lug‘atlarda lingvistik, madaniy va didaktik komponentlarni integratsiyalash ularning ta’limiy va madaniyatlararo ahamiyatini oshirishi asoslab beriladi.

KALIT SO‘ZLAR

AI-lug‘at, etnomadaniy birliklar, lingvomadaniyatshunoslik, leksikografiya, lingvodidaktika, madaniy semantika, raqamli leksikografiya, madaniyatlararo kompetensiya.

АННОТАЦИЯ

В данной статье анализируются лингвистические и лингводидактические основы интерпретации этнокультурных единиц в AI-словарях. Выявляются основные проблемы представления культурно маркированной лексики в цифровых лексикографических системах и обосновывается необходимость многомерной, культурно ориентированной модели толкования. Показано, что упрощённые и универсализированные определения не обеспечивают адекватного понимания этнокультурной семантики, и предлагается

ISSN: 2692-515-x, Impact Factor: 9.23
American Academic publishers, volume 6, issue 01,2026
Journal: https://www.academicpublishers.org/journals/index.php/ijai

система принципов её более точного описания, включая сохранение культурной маркированности, многоуровневую интерпретацию и адаптивную подачу материала. Доказывается, что интеграция лингвистического, культурного и педагогического компонентов в AI-словарях повышает их образовательную и межкультурную ценность.

КЛЮЧЕВЫЕ СЛОВА

AI-словарь, этнокультурные единицы, лингвокультурология, лексикография, лингводидактика, культурная семантика, цифровая лексикография, межкультурная компетенция.

Introduction

In recent years, artificial intelligence technologies have been rapidly penetrating various fields of humanities, including linguistics and lexicography. Traditional printed dictionaries are gradually being replaced by digital and AI-based lexicographic systems that offer not only word meanings, but also contextual examples, usage recommendations, semantic associations and multimodal support (Lew, 2012; Granger & Paquot, 2012). Such changes have significantly expanded the functional possibilities of dictionaries and turned them into complex interactive information resources.

It is well known that language is not only a means of communication, but also a repository of culture. This idea has been repeatedly emphasized in linguistic and linguocultural studies, where language is viewed as a reflection of national mentality and worldview (Sapir, 1921; Humboldt, 1988; Wierzbicka, 1997). In this respect, ethnocultural units occupy a special place in the lexical system of any language. They include words and expressions that are closely connected with traditions, customs, social institutions, historical realities and cultural symbols.

However, practice shows that many AI-based dictionaries still interpret such units in a simplified or universalized way, often ignoring or minimizing their cultural specificity (Bergenholtz & Tarp, 2010). As a result, the user receives only a partial understanding of the meaning, which may be sufficient for basic communication, but is insufficient for deep intercultural comprehension and correct pragmatic use.

Therefore, the purpose of this article is to analyze the linguistic and linguodidactic foundations of interpreting ethnocultural units in AI-based dictionaries and to justify the necessity of a more integrated, culture-oriented approach to their lexicographic description.

Ethnocultural Units as a Linguistic Phenomenon

Ethnocultural units can be defined as lexical units that reflect culturally specific concepts, objects, traditions, rituals, social relations and values characteristic of a particular linguistic community (Wierzbicka, 1997; Maslova, 2001). They often include realia, culture-specific terms and conceptually loaded expressions that have no full equivalents in other languages (Vlahov & Florin, 1980).

ISSN: 2692-515-x, Impact Factor: 9.23
American Academic publishers, volume 6, issue 01,2026
Journal: https://www.academicpublishers.org/journals/index.php/ijai

From a linguistic point of view, such units are characterized by a complex semantic structure in which denotative meaning is closely intertwined with connotative, evaluative and associative components (Apresjan, 2000). For example, words denoting national dishes, traditional clothing or social institutions usually carry additional cultural information that cannot be deduced from a simple dictionary definition.

In linguoculturology, ethnocultural units are considered as verbalized fragments of the national worldview and as carriers of cultural memory (Maslova, 2001; Karasik, 2002). Therefore, their interpretation requires going beyond purely linguistic explanation and involving cultural, historical and social commentary.

Interpretation of Ethnocultural Units in AI-Based Dictionaries: Linguistic and Technological Aspect

In contemporary digital lexicography, AI-based dictionaries are increasingly built on natural language processing (NLP) technologies, machine learning models and large language models (LLMs), which allow systems to process large corpora of texts, identify semantic patterns and generate context-sensitive explanations (Granger & Paquot, 2012; Lew, 2012). In theory, such technologies create favorable conditions for a more flexible and user-oriented representation of lexical meaning.

However, the semantic modeling used in many AI systems is still largely based on probabilistic and distributional principles. While this approach is effective for describing general vocabulary, it often proves insufficient for ethnocultural units, whose meaning is not fully derivable from contextual frequency alone (Wierzbicka, 1997). The cultural component of meaning is not merely a sum of usage contexts, but a result of historical, social and symbolic development.

Another important feature of modern AI-based dictionaries is their multimodal nature**.** They may include images, audio, video fragments and hyperlinks in addition to verbal explanations (Kress & van Leeuwen, 2001). This creates new opportunities for representing ethnocultural units, especially those connected with material culture, rituals or everyday practices.

Nevertheless, even multimodality does not automatically guarantee cultural adequacy. If the system is not conceptually oriented toward preserving cultural specificity, visual or contextual supplements may remain purely illustrative and lead to semantic flattening and cultural simplification (Bergenholtz & Tarp, 2010).

From a linguistic point of view, ethnocultural units should be treated as conceptually dense signs that require layered interpretation:

- denotative meaning,
- cultural-historical commentary,
- pragmatic usage,
- associative and symbolic value (Apresjan, 2000; Karasik, 2002).

Thus, AI-based dictionaries should move from one-dimensional definitions to multilevel semantic representation models**.**

ORIGINAL ARTICLE

INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE

ISSN: 2692-515-x, Impact Factor: 9.23

American Academic publishers, volume 6, issue 01,2026

Journal: https://www.academicpublishers.org/journals/index.php/ijai

Linguodidactic Aspect: Educational Potential of AI-Based Dictionaries

From the linguodidactic point of view, dictionaries are not only reference tools, but also important instruments of language learning and teaching (Hartmann, 2001). AI-based dictionaries significantly expand these didactic possibilities due to their interactivity, adaptability and multimodal nature.

Modern language education increasingly emphasizes the importance of intercultural communicative competence (Byram, 1997), which includes the ability to understand and interpret culture-specific concepts, values and communicative norms. Ethnocultural units play a key role in this process, because they represent one of the most culturally marked layers of vocabulary.

If AI-based dictionaries provide only simplified or culturally neutral explanations, learners may form distorted or superficial ideas about the target culture. On the contrary, if the dictionary integrates linguistic explanation with cultural and pragmatic commentary, it can become an effective tool for developing intercultural awareness and interpretative skills (Byram, 1997; Hartmann, 2001).

In practical lexicographic use, the problems of interpreting ethnocultural units in AI-based dictionaries become especially visible when concrete culture-bound items are queried by users. For example, if a learner searches for the Uzbek word *"mahalla"*, many digital systems simply explain it as "a neighborhood" or "a community", which is formally correct but culturally insufficient, because *mahalla* also denotes a socially regulated institution with specific functions, norms of behavior and collective responsibility. Similarly, the word *"palov"* is often translated as "rice dish" or "pilaf", while in Uzbek culture it is not merely a type of food but a symbolically marked element of social rituals, hospitality and family events. The same problem can be observed in English–Uzbek or English–Russian directions with words like *"Thanksgiving"*, *"pub"* or *"gentleman"*, which in many AI-generated explanations are reduced to simplified labels and lose their cultural and pragmatic depth. In such cases, the user receives an operationally useful but conceptually poor explanation, which is sufficient for basic reference but inadequate for real intercultural communication or translation. A more effective AI-based dictionary model should therefore provide layered entries in which a short functional definition is accompanied by a brief cultural note, typical situations of use and, where possible, visual or contextual illustration. This kind of practical structure would allow the user to understand not only "what the word means", but also "how and in what cultural logic it is used", which is especially important for learners, translators and users working in multicultural communicative environments.

Conceptual Principles for Representing Ethnocultural Units in AI-Based Dictionaries

To improve the interpretation of ethnocultural units, several principles should be implemented:

1. Cultural markedness preservation – explicit indication of culture-specific status (Wierzbicka, 1997).

ISSN: 2692-515-x, Impact Factor: 9.23
American Academic publishers, volume 6, issue 01,2026
Journal: https://www.academicpublishers.org/journals/index.php/ijai

2. Multilevel interpretation – combining definition, commentary, pragmatics and examples (Apresjan, 2000).
3. Adaptive presentation – different levels for learners, translators and researchers (Bergenholtz & Tarp, 2010).

Such a model transforms the dictionary into an active educational environment and supports the development of intercultural competence (Byram, 1997).

Conclusion

The analysis shows that the interpretation of ethnocultural units in AI-based dictionaries is a complex interdisciplinary problem requiring the integration of linguistic, cultural and pedagogical approaches. AI technologies should serve not only automation, but also deep semantic and cultural representation. Multilevel, culturally oriented models will significantly enhance both lexicographic quality and educational value.

References

1. Apresjan, Y. D. (2000). *Systematic Lexicography*. Oxford: Oxford University Press.
2. Bergenholtz, H., & Tarp, S. (2010). *Lexicography at a Crossroads*. Bern: Peter Lang.
3. Byram, M. (1997). *Teaching and Assessing Intercultural Communicative Competence*. Clevedon: Multilingual Matters.
4. Granger, S., & Paquot, M. (2012). Electronic lexicography. In *Oxford Handbook of Lexicography*. Oxford: OUP.
5. Hartmann, R. R. K. (2001). *Teaching and Researching Lexicography*. London: Longman.
6. Humboldt, W. von. (1988). *On Language*. Cambridge: CUP.
7. Karasik, V. I. (2002). *Language Circle: Personality, Concepts, Discourse*. Volgograd.
8. Kress, G., & van Leeuwen, T. (2001). *Multimodal Discourse*. London: Arnold.
9. Lew, R. (2012). How can we make electronic dictionaries more effective? *Lexikos*, 22.
10. Maslova, V. A. (2001). *Linguoculturology*. Moscow: Academia.
11. Sapir, E. (1921). *Language*. New York: Harcourt.
12. Vlahov, S., & Florin, S. (1980). *Untranslatable in Translation*. Moscow.
13. Wierzbicka, A. (1997). *Understanding Cultures through Their Key Words*. Oxford: OUP.