ISSN:2828-6235
Impact factor:13.77

International Journal of Literature and Language Studies(IJLLS)

SEMANTIC TRANSFORMATION MODELS OF CULTURE-SPECIFIC REALIA IN GENERATIVE AI TRANSLATION BETWEEN ENGLISH AND UZBEK

Mardayev Omon To‘lqinovich
PhD researcher, Tashkent State University of Oriental Studies
doi:https://doi.org/10.5281/zenodo.22988483

Abstract: This article examines semantic transformation models of culture-specific realia in generative AI translation between English and Uzbek. Particular attention is paid to the preservation, modification, generalization, explicitation, substitution, and reduction of culturally significant semantic components. The research applies comparative, contextual, componential, semantic, and pragmatic methods of analysis. The findings demonstrate that generative AI translation does not simply reproduce culture-specific meaning in another linguistic code; rather, it reconstructs such meaning through a variety of semantic operations. Six principal transformation models are identified: direct cultural transfer, semantic approximation, semantic generalization, explicitation, cultural substitution, and semantic-pragmatic reduction. The analysis indicates that linguistically fluent AI-generated translations may still demonstrate partial cultural loss, especially when complex Uzbek concepts such as mahalla, hashar, dasturxon, to‘r, choyxona, and kelin salom are reduced to more general English equivalents. The article argues that the evaluation of generative AI translation should therefore incorporate cultural-semantic adequacy alongside grammatical accuracy, lexical correspondence, and textual fluency.

Keywords: generative artificial intelligence, culture-specific realia, English and Uzbek, semantic transformation, AI translation, cultural semantics, pragmatic adaptation, equivalence, explicitation, cultural substitution.

Introduction: The development of digital technologies has brought about profound changes in the theory and practice of translation. Whereas the first generations of machine translation relied mainly on predefined linguistic rules and bilingual correspondences, later statistical systems began to calculate translation probabilities on the basis of large collections of parallel texts. Neural machine translation introduced a further qualitative change by enabling computational systems to process linguistic units within wider contextual environments. The emergence of generative artificial intelligence and large language models has expanded these possibilities even further. Modern generative systems can produce coherent target texts, take account of contextual relationships, modify stylistic registers, explain unfamiliar concepts, and select among alternative translation solutions according to the instructions provided by the user.

However, high linguistic fluency does not necessarily mean that the cultural content of the source text has been fully preserved. This distinction becomes particularly important when the source text contains culture-specific realia. Realia represent lexical units denoting objects, phenomena, traditions, social institutions, customs, ceremonies, food, clothing, forms of interaction, and other elements that are closely associated with the historical and cultural experience of a particular linguistic community. Their meanings are not limited to dictionary definitions. They often contain additional semantic features connected with collective memory, social values, historical experience, symbolic associations, and culturally determined patterns of behaviour.

The translation of such units has traditionally been regarded as one of the most complicated issues in translation studies. Nida (1964) demonstrated that translation equivalence cannot always be established through formal correspondence and emphasized the importance of the effect produced on the target-language recipient. Newmark (1988), examining cultural words and translation procedures, proposed strategies such as transference, descriptive equivalent, functional

equivalent, cultural equivalent, and componential analysis. Baker (2018) likewise emphasizes that non-equivalence at word level may arise when the source-language concept is not lexicalized in the target language or when culturally specific distinctions are encoded differently. Venuti's (2018) concepts of domestication and foreignization further demonstrate that translation involves a strategic decision concerning whether cultural difference should be preserved or adapted to the norms of the receiving culture.

The relevance of these theoretical approaches becomes even greater in generative AI translation. In human translation, the translator consciously interprets a culturally marked unit, identifies its communicative function, considers the assumed knowledge of the target reader, and chooses an appropriate translation procedure. Generative AI performs a comparable operation computationally on the basis of linguistic patterns, contextual probability, training data, and user instructions. The resulting translation may be highly fluent, but the semantic transformation that occurs between the source cultural concept and its target representation requires separate investigation.

This problem is especially relevant to the English–Uzbek language pair. English and Uzbek differ not only structurally but also in the cultural environments represented through their lexical systems. Uzbek contains numerous words that reflect traditional forms of community organization, hospitality, family relations, ceremonies, material culture, national cuisine, and social etiquette. Units such as mahalla, hashar, dasturxon, to‘r, sumalak, choyxona, do‘ppi, kelin salom, oqsoqol, and to‘y cannot always be replaced by single English words without some degree of semantic modification.

For instance, translating mahalla as neighborhood may be acceptable in a context where only a residential area is relevant. Yet mahalla in Uzbek cultural discourse can simultaneously refer to a territorial unit, a community, a network of social relationships, a local institution, and a culturally significant form of collective identity. Consequently, the English word neighborhood may reproduce the spatial component while leaving institutional and sociocultural meanings underrepresented.

A similar phenomenon can be observed in the translation of dasturxon. Depending on context, this word can refer to a cloth on which food is served, the food placed before guests, a shared meal, or, metaphorically, a culturally significant space of hospitality and social interaction. If generative AI translates dasturxon simply as tablecloth, the physical component of the meaning may be preserved while the culturally important associations of hospitality, generosity, family gathering, and respect for guests disappear.

Therefore, the main problem addressed in this research concerns not merely whether generative AI can translate culture-specific realia, but how their semantic structure changes in the process of AI-mediated translation. This distinction is important because an AI-generated translation may be grammatically correct and stylistically natural while still being culturally incomplete.

The aim of the study is to investigate the principal semantic transformation models through which culture-specific realia are represented in generative AI translation between English and Uzbek and to determine how such transformations influence cultural-semantic adequacy.

The study also seeks to demonstrate that AI translation should be understood as a process of semantic reconstruction. The source-language cultural unit enters the translation process with a particular combination of denotative, connotative, pragmatic, historical, and associative features. During translation, some of these features may be retained, others may be reformulated, and still others may disappear. The target-language expression is consequently not always an identical semantic copy of the original but a reconstructed representation shaped by the resources and cultural conventions of the target language.

Materials and Methods: The research was conducted within a qualitative comparative framework. The principal material consists of culture-specific lexical units and expressions characteristic of English and Uzbek cultural discourse and their possible representations in generative AI-mediated translation. The selection of material was based on the principle of cultural markedness. In other words, priority was given to units whose meanings are closely associated with a specific cultural practice, institution, tradition, object, ceremony, or pattern of social interaction and which do not always possess a complete lexical equivalent in the other language.

The Uzbek material includes lexical units related to traditional social organization, hospitality, ceremonies, national cuisine, clothing, and everyday cultural practices. Particular attention is given to such items as mahalla, hashar, dasturxon, to‘r, choyxona, sumalak, do‘ppi, kelin salom, and oqsoqol. These units were selected because their semantic structures extend beyond simple referential meanings. They contain cultural information that may not be immediately accessible to English-speaking recipients.

English culture-specific units can create comparable problems when translated into Uzbek. Expressions and concepts such as Thanksgiving, Boxing Day, pub, afternoon tea, prom, county, and Sunday roast are embedded in specific historical and social contexts. A literal translation may communicate the general meaning of these units but fail to reproduce their cultural associations. For example, prom cannot always be adequately represented by a general Uzbek equivalent meaning “school party”, since the English-language cultural concept includes associations with a formal event for secondary-school students, particular forms of dress, graduation-related traditions, and specific social practices.

Several interconnected analytical methods were employed. Comparative analysis was used to determine the relationship between the semantic structures of source and target units. Componential analysis made it possible to divide culturally marked meanings into individual semantic components and to identify which components were retained or eliminated in translation. Contextual analysis was applied because the meaning of realia frequently depends on the communicative situation. The same cultural word can require different translation strategies in literary, academic, tourist, journalistic, or everyday contexts.

Pragmatic analysis was also important because the cultural function of a lexical unit cannot be reduced to its referential content. A translation may communicate what an object is while failing to communicate what that object means within the source culture. Thus, the analysis considers both denotative correspondence and the preservation of cultural-pragmatic meaning.

The research procedure was organized around three basic analytical stages. First, the semantic and cultural structure of the source-language unit was identified. Second, the target-language representation was examined in terms of lexical and contextual correspondence. Third, the semantic change between the two representations was classified according to the type of transformation involved. This procedure enabled the study to distinguish between preservation, approximation, generalization, explanation, substitution, and reduction.

Results: The analysis demonstrates that culture-specific realia do not follow a single translation pattern in generative AI environments. Instead, their target-language representations display different degrees of semantic preservation and cultural adaptation. Six recurrent transformation models can be distinguished.

3.1. Direct Cultural Transfer

Direct cultural transfer occurs when the original culturally marked unit is retained in the target-language text, usually through transliteration or transcription. This strategy is particularly appropriate when the target language lacks an established equivalent and when preservation of cultural identity is important.

For example, in the sentence Navro‘z bayramida sumalak tayyorlanadi, the culturally important lexical unit sumalak may be preserved in English: Sumalak is prepared during the

Navruz celebrations. Such a translation maintains the cultural identity of the dish instead of replacing it with a generalized expression such as traditional food or wheat dish.

The advantage of direct transfer is that it minimizes cultural substitution. The target reader encounters the source-cultural concept as an independent cultural unit. At the same time, direct transfer alone may not guarantee comprehension. A reader unfamiliar with Uzbek culture may not know what sumalak means. Consequently, direct transfer frequently functions most effectively when combined with contextual explanation.

The same principle applies to do‘ppi. Rendering the word merely as cap or hat communicates its general physical function but removes its cultural specificity. Retaining doppi introduces the Uzbek cultural object into the English text. In academic, ethnographic, historical, and cultural texts, this strategy can be particularly valuable because the purpose of translation is not simply to make the text immediately familiar but also to preserve information about the source culture.

3.2. Semantic Approximation

Semantic approximation is observed when generative AI chooses a target-language lexical item that shares the central denotative meaning of the source realia but does not reproduce its complete cultural-semantic structure. This model is particularly common when the target language contains a concept that appears functionally similar to the source unit.

The translation of mahalla as neighborhood provides a clear example. At first glance, both terms refer to a geographically localized residential community. However, their semantic structures are not identical. In Uzbek cultural and social practice, mahalla may include territorial, institutional, interpersonal, administrative, moral, and symbolic dimensions. It can function as a form of community organization and as an important environment for collective identity and social interaction.

The English neighborhood, by contrast, primarily refers to a geographical area and the people who live in it. Consequently, the translation preserves the spatial and community components but may reduce the institutional and ethnocultural aspects.

This example demonstrates an important feature of generative AI translation: lexical plausibility does not necessarily indicate complete semantic equivalence. An output may sound entirely natural to an English-speaking reader while representing only part of the source concept.

The same issue occurs in the translation of oqsoqol as elder. The English word communicates age and, in certain contexts, authority or respect. The Uzbek concept, however, may additionally imply recognized social status, traditional authority, community leadership, and culturally institutionalized respect. Thus, elder can function as an approximate equivalent, but its semantic scope is not necessarily identical to that of oqsoqol.

3.3. Semantic Generalization

Semantic generalization occurs when a specific culturally marked unit is translated by a broader and culturally less specific target-language expression. Generative AI may employ this transformation in order to produce a fluent and easily understandable target text.

For instance, do‘ppi may be translated as traditional hat. Such a translation is understandable, but the semantic structure becomes more general. The source unit refers to a particular form of traditional headwear with culturally recognizable shapes, patterns, regional variations, and symbolic associations. The target expression traditional hat communicates only the broad category.

A similar process can occur with sumalak, which may become traditional Uzbek dish, or hashar, which may be generalized as community work. In each case, the target-language expression conveys the general conceptual category while reducing the distinctive cultural characteristics of the source unit.

Generalization therefore represents a compromise between accessibility and cultural precision. It may be appropriate when detailed cultural information is not central to the communicative

purpose of the text. However, in ethnographic, literary, historical, or linguocultural translation, excessive generalization can lead to substantial semantic loss.

3.4. Explicitation as Cultural-Semantic Reconstruction

Explicitation represents one of the most productive ways of dealing with culturally asymmetric concepts. In this transformation, generative AI preserves or translates the source realia while adding information necessary for the target reader to understand its cultural function.

Consider the sentence Yakshanba kuni mahallada hashar o'tkazildi. Translating this merely as Community work was held in the neighborhood on Sunday communicates the practical activity but eliminates the culturally specific status of hashar. A more culturally informative version would be: On Sunday, the mahalla organized a hashar, a traditional form of voluntary collective community work.

In this version, both mahalla and hashar remain visible as source-cultural concepts, while explanatory information makes them comprehensible to an English-speaking reader. The translation consequently performs two functions simultaneously: preservation of cultural identity and transmission of semantic information.

The same method is effective for kelin salom. A literal translation such as bride's greeting communicates only the lexical surface. In Uzbek culture, however, kelin salom refers to a traditional ceremonial practice in which the bride formally greets members of the groom's family and relatives. Therefore, an explanatory translation such as kelin salom, a traditional Uzbek ceremony in which the bride formally greets the groom's relatives provides a considerably higher degree of cultural-semantic adequacy.

Explicitation is especially significant in generative AI translation because large language models can potentially produce contextual explanations without requiring a fixed dictionary equivalent. This represents one of the important advantages of generative systems over earlier word-based approaches. Nevertheless, the quality of such explanation depends on whether the AI system has accurately interpreted the source concept. Incorrect or overly generalized explanations can themselves become a source of cultural distortion.

3.5. Cultural Substitution

Cultural substitution occurs when a source-cultural unit is replaced with a concept that is more familiar to the target-language recipient. This strategy can increase readability but also presents a considerable risk of cultural displacement.

The translation of choyxona as café illustrates this problem. Both places may serve drinks and food and function as spaces for social interaction. Yet their cultural structures differ. A traditional choyxona is associated with Central Asian tea culture, forms of communal interaction, traditional cuisine, architectural organization, and historically established social practices. A café belongs to a broader international or Western-oriented cultural model.

If choyxona is systematically translated as café, the target reader receives an easily understandable image but may interpret the Uzbek cultural environment through an inappropriate target-cultural category. Thus, communicative accessibility is achieved at the cost of cultural specificity.

A similar problem arises when to‘y is automatically rendered as wedding. In many contexts this is correct because to‘y can indeed refer to a wedding celebration. However, Uzbek usage of to‘y is semantically broader and may refer to different types of ceremonial celebrations. Context must therefore determine whether wedding, celebration, ceremony, or a more specific explanatory equivalent is appropriate.

This demonstrates that generative AI must interpret not only lexical meanings but also contextual cultural functions. When it chooses a culturally familiar substitute without considering these functions, the resulting translation may represent a form of semantic domestication.

3.6. Semantic-Pragmatic Reduction

Semantic-pragmatic reduction constitutes one of the most significant risks in AI translation of realia. It occurs when the central referential meaning is preserved but culturally important connotations, associations, social values, or pragmatic functions are lost.

The Uzbek word dasturxon provides a particularly revealing example. At the material level, it may denote the cloth or surface on which food is served. In actual cultural discourse, however, its meaning is much wider. Dasturxon can evoke hospitality, family unity, generosity, respect for guests, shared meals, abundance, and social closeness. Expressions connected with an ochiq dasturxon or preparing a dasturxon for guests cannot always be interpreted solely in terms of a physical tablecloth.

When AI chooses tablecloth, it may therefore preserve the object while reducing the cultural concept. When it chooses meal or table, another part of the meaning may be highlighted, but other semantic components remain absent. The most appropriate translation depends on context.

Another illustrative example is to‘r. In traditional Uzbek spatial and hospitality practices, to‘r does not simply refer to a particular part of a room. It is associated with status, honour, and respect. Inviting a guest to the to‘r functions as a culturally meaningful act of hospitality. If the sentence Mehmonni to‘rga taklif qilishdi is translated simply as They asked the guest to sit down, the general communicative event is conveyed but the culturally significant relationship between spatial position and social respect disappears.

A culturally sensitive translation could instead state that the guest was invited to sit in the place of honour. Such a solution demonstrates that successful translation requires identification of the pragmatic meaning encoded in the source cultural practice.

Discussion: The findings indicate that generative AI translation of culture-specific realia should be viewed as a process of semantic negotiation between two linguistic and cultural systems. The source-language unit carries a set of culturally determined semantic components, whereas the target language offers lexical and conceptual resources that may only partially correspond to them. Generative AI must therefore select which components to preserve, which to explain, and which to adapt.

One of the most important observations emerging from the analysis is that translation fluency and cultural adequacy are separate parameters. Modern generative AI can produce highly natural English or Uzbek sentences even when the culturally marked meaning has been simplified. This characteristic can make semantic loss difficult to detect because a fluent translation often creates an impression of accuracy.

For example, translating mahalla raisi as neighborhood leader may sound natural, but it can reduce the institutional meaning associated with the Uzbek mahalla system. Similarly, rendering choyxona as café or dasturxon as tablecloth may result in grammatically perfect sentences while transforming the cultural worldview represented by the original text.

This observation corresponds with broader theories of translation equivalence. Nida's concept of dynamic equivalence emphasizes communicative effect, while Newmark's translation procedures demonstrate that different cultural words require different solutions. In the context of generative AI, however, these procedures are not necessarily selected through conscious theoretical reasoning. Rather, they emerge from probabilistic language generation influenced by context, training data, and prompt formulation.

Venuti's distinction between domestication and foreignization is particularly useful for understanding this process. Generative AI tends toward domestication when it chooses familiar

target-language concepts such as neighborhood, café, or traditional hat. It moves toward foreignization when it retains mahalla, choyxona, or do‘ppi and supplies contextual explanation.

Neither approach is universally correct. The appropriate strategy depends on the genre, communicative purpose, readership, and cultural significance of the source unit. In a short tourist information text, traditional Uzbek cap may be sufficient for do‘ppi. In an ethnographic study, however, preserving do‘ppi as a cultural term and explaining its significance would be more appropriate.

The analysis further suggests that generative AI has created a new possibility between strict foreignization and domestication: interactive cultural explicitation. Because users can instruct an AI system to preserve cultural terms, explain them, or adapt translations for particular audiences, the translation strategy can be modified through prompt design. Thus, translation becomes partly dependent on human–AI interaction.

A generic prompt requesting only translation may encourage a highly fluent target text. By contrast, an instruction to “preserve culturally specific Uzbek concepts and provide concise explanations where no complete English equivalent exists” may increase the visibility of the source culture. This implies that prompt formulation should be considered an additional variable in research on AI-mediated translation.

From a semantic perspective, the results allow culture-specific translation to be represented as a continuum. At one end lies near-complete preservation, achieved through direct transfer and appropriate explicitation. At the other end lies substantial cultural-semantic loss, which may occur through excessive generalization or reduction. Semantic approximation and cultural substitution occupy intermediate positions.

An important theoretical implication is that semantic transformation itself should not automatically be treated as a translation error. Translation between culturally distant languages inevitably requires transformation. The decisive question is whether the transformation is justified by the communicative context and whether culturally relevant information is preserved to an appropriate degree. For example, translating hashar as community work is not inherently incorrect. In a context where only the practical activity matters, such generalization may be sufficient. However, in a study of Uzbek traditions, the same translation would be inadequate because the culturally institutionalized and collective nature of hashar constitutes essential information.

Therefore, AI-generated translations should be evaluated according to the relationship between semantic transformation and communicative purpose rather than according to lexical correspondence alone.

Another important issue concerns the direction of translation. Uzbek-to-English translation may involve the neutralization of Uzbek cultural units because English lacks corresponding lexicalized concepts. Conversely, English-to-Uzbek translation may require explanations of culturally specific institutions and traditions such as Thanksgiving, Boxing Day, prom, or pub. Consequently, cultural asymmetry operates in both directions, although the nature and frequency of transformation may differ.

The findings suggest that an adequate evaluation framework for generative AI translation should include at least four interconnected criteria: linguistic accuracy, semantic adequacy, cultural preservation, and pragmatic appropriateness. A translation that performs well only in terms of grammatical correctness cannot automatically be classified as fully adequate.

Particularly in literary, ethnographic, historical, tourism, media, and intercultural communication texts, cultural preservation should constitute an independent quality criterion. The target text should allow readers not only to understand the immediate event but also, where relevant, to recognize the cultural system within which that event acquires meaning.

Conclusion: The analysis of culture-specific realia demonstrates that generative AI translation between English and Uzbek represents a complex process of semantic and cultural

transformation. The difficulties involved arise not only from structural differences between the languages but also from differences between the cultural concepts encoded in their lexical systems.

The study identified six major models through which culture-specific meaning can be transformed: direct cultural transfer, semantic approximation, semantic generalization, explicitation, cultural substitution, and semantic-pragmatic reduction. These models differ according to the degree to which they preserve the cultural information of the source unit.

Direct transfer is effective in maintaining cultural identity but may require additional explanation. Semantic approximation provides natural target-language equivalents but often preserves only the central components of culturally complex meanings.

Generalization improves comprehensibility while reducing specificity. Explicitation appears particularly productive because it enables the translator or AI system to retain the original concept while providing the target reader with sufficient cultural information. Cultural substitution may facilitate communication but risks replacing the conceptual framework of the source culture with that of the target culture. Semantic-pragmatic reduction represents the most problematic transformation when socially and culturally meaningful associations disappear from the target text.

The analysis of such Uzbek concepts as mahalla, hashar, dasturxon, to‘r, choyxona, do‘ppi, sumalak, and kelin salom demonstrates that their meanings cannot always be adequately represented through isolated English lexical equivalents. Their translation requires consideration of historical, social, pragmatic, and cultural components. The same principle applies to English realia when they are transferred into Uzbek.

References

1. Baker, M. (2018). In Other Words: A Coursebook on Translation. 3rd ed. London and New York: Routledge.
2. Bassnett, S. (2014). Translation Studies. 4th ed. London and New York: Routledge.
3. Catford, J. C. (1965). A Linguistic Theory of Translation: An Essay in Applied Linguistics. London: Oxford University Press.
4. House, J. (2015). Translation Quality Assessment: Past and Present. London: Routledge.
5. Koehn, P. (2020). Neural Machine Translation. Cambridge: Cambridge University Press.
6. Newmark, P. (1988). A Textbook of Translation. New York: Prentice Hall.
7. Nida, E. A. (1964). Toward a Science of Translating. Leiden: E. J. Brill.
8. Nida, E. A., & Taber, C. R. (1969). The Theory and Practice of Translation. Leiden: E. J. Brill.
9. O’Brien, S. (2012). Translation as human-computer interaction. Translation Spaces, 1, 101–122.
10. Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. Advances in Neural Information Processing Systems, 27, 3104–3112.
11. Toral, A., & Way, A. (2018). What level of quality can neural machine translation attain on literary text? In J. Moorkens et al. (Eds.), Translation Quality Assessment: From Principles to Practice. Cham: Springer, 263–287.
12. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.
13. Venuti, L. (2018). The Translator’s Invisibility: A History of Translation. 3rd ed. London and New York: Routledge.
14. Vinay, J.-P., & Darbelnet, J. (1995). Comparative Stylistics of French and English: A Methodology for Translation. Amsterdam: John Benjamins.

15. Vlahov, S., & Florin, S. (1980). Neperevodimoe v perevode. Moscow: Mezhdunarodnye otnosheniya.