BG Pulse Journal USA
ISSN:1947-2536 http://www.bgpulseusa.com/index.php/bg
BG PULSE JOURNAL

POST-EDITING (MTPE) PROBLEMS AND THEIR LINGUISTIC SOLUTIONS IN RENDERING UZBEK NATIONAL REALIA INTO ENGLISH THROUGH ARTIFICIAL INTELLIGENCE AND NEURAL MACHINE TRANSLATION

Azimova Sitora Orif kizi

Sirdaryo region, Sirdaryo district school N:6, English teacher

e-mail: ibrohimovas69@gmail.com

Doi: https://doi.org/10.5281/zenodo.22053093

Abstract. The article examines the problems of machine translation post-editing (MTPE) that arise when Uzbek national realia are rendered into English by neural machine translation engines and large language models. A working classification of Uzbek realia relevant to machine processing is proposed, and a typology of nine recurrent machine errors is established and mapped onto the categories of the MQM framework. The levels of post-editing intervention are analysed in accordance with ISO 18587:2017, and it is shown that the light post-editing regime is structurally incapable of resolving realia-related errors, since these belong to the cultural rather than the grammatical layer of the text. A decision procedure comprising seven linguistic solutions is proposed, together with preventive measures at the pre-editing, terminology-management and prompt-design stages. Empirical evidence from comparative evaluations of generative and neural systems on culture-bound literary material is discussed in support of a hybrid workflow in which the machine produces the draft, a controlled terminology base secures consistency, and the human post-editor retains responsibility for cultural adequacy.

Key words: post-editing, MTPE, neural machine translation, large language models, realia, culture-bound units, ISO 18587, MQM, terminology base, pre-editing, Uzbek language.

O‘ZBEK MILLIY REALIYALARINI SUN’IY INTELLEKT VA NEYRON MASHINA TARJIMASI ORQALI INGLIZ TILIGA O‘GIRISHDA POST-TAHRIR (MTPE) MUAMMOLARI VA ULARNING LINGVISTIK YECHIMLARI

Annotatsiya. Maqolada o‘zbek milliy realiyalarini neyron mashina tarjimasi tizimlari va katta til modellari yordamida ingliz tiliga o‘girishda yuzaga keladigan post-tahrir (MTPE) muammolari tahlil qilinadi. Mashinaviy qayta ishlash nuqtai nazaridan o‘zbek realiyalarining ishchi tasnifi taklif etiladi, muntazam takrorlanuvchi to‘qqizta xato turi aniqlanib, ular MQM tizimi toifalariga moslashtiriladi. Post-tahrir darajalari ISO 18587:2017 standarti asosida ko‘rib chiqilib, yengil post-tahrir rejimi realiya bilan bog‘liq xatolarni bartaraf eta olmasligi asoslanadi, chunki bu xatolar matnning grammatik emas, madaniy qatlamiga tegishlidir. Yettita lingvistik yechimdan iborat qaror qabul qilish tartibi hamda oldindan tahrirlash, terminologiyani boshqarish va so‘rov (prompt) loyihalash bosqichlaridagi profilaktik choralar taklif qilinadi.

Kalit so‘zlar: post-tahrir, MTPE, neyron mashina tarjimasi, katta til modellari, realiya, milliy-madaniy birliklar, ISO 18587, MQM, terminologik baza, oldindan tahrirlash, o‘zbek tili.

BG Pulse Journal USA
ISSN:1947-2536 http://www.bgpulseusa.com/index.php/bg
BG PULSE JOURNAL

ПРОБЛЕМЫ ПОСТРЕДАКТИРОВАНИЯ (MTPE) ПРИ ПЕРЕДАЧЕ УЗБЕКСКИХ НАЦИОНАЛЬНЫХ РЕАЛИЙ НА АНГЛИЙСКИЙ ЯЗЫК СРЕДСТВАМИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И НЕЙРОННОГО МАШИННОГО ПЕРЕВОДА

Аннотация. В статье рассматриваются проблемы постредактирования машинного перевода (MTPE), возникающие при передаче узбекских национальных реалий на английский язык системами нейронного машинного перевода и большими языковыми моделями. Предложена рабочая классификация узбекских реалий с точки зрения машинной обработки, установлена типология девяти регулярных машинных ошибок и произведено их соотнесение с категориями системы MQM. Уровни постредактирования проанализированы в соответствии со стандартом ISO 18587:2017; показано, что режим лёгкого постредактирования принципиально не способен устранить ошибки, связанные с реалиями, поскольку они относятся к культурному, а не к грамматическому слою текста. Предложена процедура принятия решений, включающая семь лингвистических решений, а также превентивные меры на этапах предредактирования, управления терминологией и проектирования запроса.

Ключевые слова: постредактирование, MTPE, нейронный машинный перевод, большие языковые модели, реалии, культурно-маркированные единицы, ISO 18587, MQM, терминологическая база, предредактирование, узбекский язык.

1. INTRODUCTION

Over the past decade the translation industry has moved from post-editing as an occasional auxiliary practice to post-editing as the default production model. Neural machine translation has raised baseline fluency to a level at which raw output is frequently indistinguishable from human writing at the sentence level, and the appearance of large language models has added a further capability — the ability to follow instructions and to take context into account. In this configuration the translator's role shifts from producing text to evaluating and correcting it, and the quality of the final product depends less on the engine than on the competence of the post-editor and on the design of the workflow.

This shift has been formalised in ISO 18587:2017, which sets out the requirements for the post-editing of machine translation output. The standard distinguishes two regimes. Light post-editing is confined to correcting errors that impede comprehension, leaving the text merely understandable. Full post-editing requires accuracy, terminological consistency, grammatical correctness and a style comparable to that of human translation, with the result being fit for purpose and equivalent in quality to a professionally translated product. The standard further requires that post-editors possess documented linguistic qualifications, translation experience and relevant subject-matter expertise.

The Uzbek–English direction, however, presents conditions in which these general provisions require substantial refinement. Uzbek remains a comparatively low-resource language for machine translation: parallel corpora are limited, the domain coverage is uneven, and the orthographic situation is complicated by the parallel circulation of Latin and Cyrillic scripts and by the unstable representation of the modifier letters o‘ and g‘. On top of these systemic difficulties lies the layer that constitutes the object of the present article — the national realia, the culture-bound units that

have no ready equivalent in the target language and whose correct rendering demands not linguistic competence alone but cultural knowledge.

The relevance of the problem is not merely theoretical. Uzbek cultural, tourist, literary and media content is increasingly localised into English through machine-assisted workflows, and the realia are precisely the elements that carry the national specificity such content is intended to convey. An error in a grammatical form damages the surface of the text; an error in a realia unit damages its purpose.

2. MATERIAL AND METHODS

The material of the study comprises Uzbek culture-bound units of the ethnographic, gastronomic, social, folkloric and etiquette spheres together with their English renderings produced by neural machine translation engines and by generative language models. The analysis is qualitative and comparative. It employs componential analysis of meaning, the classification of realia in the tradition of Vlakhov and Florin, the inventory of translation procedures proposed by Newmark and by Vinay and Darbelnet, the taxonomy of extralinguistic cultural references of Pedersen, the error categories of the MQM framework, and the post-editing requirements of ISO 18587:2017. The notion of post-editing effort is used in the tripartite division into temporal, technical and cognitive effort.

3. UZBEK REALIA AS AN OBJECT OF MACHINE PROCESSING

For the purposes of machine translation, the conventional classification of realia has to be supplemented by a parameter that is irrelevant to human translation but decisive for automatic processing: the degree to which the unit is represented in the training data. A realia unit that occurs frequently in parallel corpora — typically because it has entered international usage — behaves in machine output quite differently from a unit confined to intra-cultural discourse (Table 1).

Table 1

Classification of Uzbek realia from the standpoint of machine processing

Group	Examples	Corpus representation	Typical machine behaviour
Gastronomic	osh, palov, non, somsa, sumalak	Partial (palov, somsa)	Substitution by an approximate analogue
Ethnographic and material	do‘ppi, chapan, beshik, dasturxon, ayvon	Low	Generalisation or literal misreading
Social and institutional	mahalla, oqsoqol, hashar, kelin	Very low	Unstable variation across the text
Ritual and calendar	Navro‘z, beshik to‘yi,	Partial (Navro‘z)	Omission of the

Group	Examples	Corpus representation	Typical machine behaviour
	sunnat to‘yi		modifier
Folkloric and onomastic	Alpomish, Go‘ro‘g‘li, Nasriddin Afandi	Low	Transliteration without a classifier
Etiquette and address	aka, opa, -jon, assalomu alaykum	Formulaic only	Flattening or complete omission

The table reveals an asymmetry that has direct methodological consequences. Units with partial corpus representation are the most dangerous ones, not the least represented ones. A unit that the engine does not recognise at all is usually transliterated, and the transliteration signals to the reader that a foreign concept is involved; a unit that the engine recognises approximately is confidently replaced by a near-analogue, and the reader receives no signal that anything has been lost. The confident error is therefore more damaging than the visible gap.

4. A TYPOLOGY OF MACHINE ERRORS IN RENDERING REALIA

Comparison of machine output with reference renderings allows nine recurrent error types to be distinguished. Their correlation with the top-level categories of the MQM framework — terminology, accuracy, linguistic conventions, style, locale conventions, audience appropriateness and design — is given in Table 2.

Table 2

Typology of machine errors in rendering Uzbek realia

Error type	Mechanism and illustration	MQM category	Detectability
Over-domestication	Replacement by a target-culture analogue: palov → risotto	Accuracy	Low
Unmotivated generalisation	do‘ppi → cap; the ethnic marker disappears	Accuracy	Low
Polysemy resolution error	osh (festive dish / wedding feast) → soup	Accuracy	Medium
Terminological inconsistency	mahalla rendered differently in successive sentences	Terminology	High

Error type	Mechanism and illustration	MQM category	Detectability
Orthographic instability	o‘ / g‘ lost or replaced; do‘ppi → duppi / dopi	Linguistic conventions	High
Modifier omission	beshik to‘yi → wedding; the specifying element is dropped	Accuracy	Medium
Hallucinated explication	Insertion of an invented descriptive detail	Accuracy	Low
Register flattening	aka, -jon neutralised or deleted	Style	Medium
Polarity reversal in paremias	The pragmatic vector of a proverb is inverted	Accuracy	Low

The right-hand column of the table is of particular practical importance. Errors of high detectability — orthographic instability and terminological inconsistency — are formal, mechanical and amenable to automated quality assurance. Errors of low detectability — over-domestication, hallucinated explication and polarity reversal — are invisible in the target text alone: the English sentence is fluent, idiomatic and internally coherent, and only comparison with the source reveals the distortion. It follows that a post-editing procedure conducted without systematic consultation of the source text is structurally unable to catch precisely the error class that matters most for realia.

Comparative evaluations of translation systems on culture-bound literary material confirm this picture from another angle. In a study of Uyghur literary translation, conventional neural engines were found to produce not only a high volume of formal errors — an average of 31.6 errors per chapter for one of the systems, of which 23.9 were spelling errors — but also polarity reversals in ethically charged maxims, whereby an imperative to be careful was rendered as its opposite. A generative model with a concise instruction produced an average of 1.4 errors and was rated by professional translators at 4.55 out of 5 for cultural fidelity, against 1.80 for the weakest neural engine. Two observations follow. First, generative models handle culture-bound material appreciably better than earlier neural engines. Second, the very fluency that makes them better also makes their residual errors harder to detect, which raises rather than lowers the demands placed on the post-editor.

5. LEVELS OF POST-EDITING AND THE REALIA PROBLEM

The distinction between light and full post-editing drawn in ISO 18587:2017 acquires a specific content when applied to culture-bound units (Table 3).

Table 3

Post-editing regimes and their applicability to realia

Parameter	Light post-editing	Full post-editing	Realia requirement
Objective	Comprehensible text	Quality comparable to human translation	Full post-editing
Errors corrected	Those impeding comprehension	All meaning-affecting errors	Full post-editing
Terminology	Not systematically checked	Aligned with approved glossaries	Full post-editing
Style and register	Not addressed	Brought into conformity	Full post-editing
Source consultation	Selective	Systematic	Systematic

The conclusion is unambiguous: light post-editing cannot in principle resolve the realia problem. Its defining criterion is comprehensibility, and a text in which palov has become risotto is entirely comprehensible. The error lies not in the grammatical but in the cultural layer, and this layer is outside the scope of the light regime by definition. Any project whose purpose is to convey national specificity must therefore be planned as a full post-editing project from the outset; the attempt to economise by selecting the light regime does not reduce cost but relocates the loss into the quality of the cultural content.

The structure of post-editing effort is affected accordingly. Of its three components — temporal, technical and cognitive — realia units load the cognitive component above all. The technical operation of replacing one lexical item with another takes seconds; reaching the decision as to which of several admissible procedures is appropriate in the given context may take considerably longer than translating the sentence from scratch. This is the source of a well-known paradox of post-editing: on culturally saturated material, correcting the machine can be slower than translating without it.

6. LINGUISTIC SOLUTIONS: A DECISION PROCEDURE

The post-editor's repertoire in dealing with realia comprises seven procedures. They are not interchangeable, and the choice among them is governed by the function of the unit in the given text rather than by the preference of the editor (Table 4).

Table 4

Linguistic solutions and the conditions of their application

Solution	Content	Condition of use	Principal risk
Retention	The Uzbek form is preserved in standardised transliteration	The unit is thematised in the text	Opacity for the reader
Retention with a classifier	The form plus a one- or two-word gloss: do‘ppi, an embroidered skullcap	First occurrence in the text	Expansion of length
Descriptive translation	Explanatory paraphrase: mahalla, a traditional neighbourhood community	The unit is central to the content	Loss of economy
Functional analogue	A target-culture unit of comparable function	Cultural specificity is not thematised	Cultural substitution
Couplet	Two procedures combined: transliteration plus generic term	Terminologically dense passages	Redundancy
Compensation	The lost element is restored elsewhere	Paremias, wordplay, register markers	Shift of equivalence
Established equivalent	A rendering already fixed in usage	The unit is codified in reference sources	Obsolescence of the source

The procedures may be ordered into a decision sequence. The first question concerns the status of the unit: if an established equivalent exists in authoritative sources, it is to be used, and the machine output is corrected against it. If it does not, the second question concerns the function of the unit in the text: if the national specificity is itself the subject of the passage, the functional analogue is excluded and retention with a classifier or descriptive translation is required. If the unit is peripheral and the visual or situational context supports it, generalisation or the functional analogue is admissible. The third question concerns position: an explanatory gloss is supplied at the first occurrence, after which the bare transliterated form is used, which preserves both intelligibility and economy.

A separate observation concerns the etiquette layer. Address forms such as aka, opa and the suffix -jon are systematically flattened by machine systems, since they possess no lexical equivalent in English and their function is relational rather than referential. Compensation is the appropriate procedure here: the deference or affection lost with the address form is redistributed across the syntax and lexical choice of the surrounding utterance. This is a solution that no engine performs of its own accord, since it requires operating at the level of the passage rather than the sentence.

7. PREVENTIVE MEASURES: PRE-EDITING, TERMINOLOGY AND PROMPT DESIGN

Post-editing corrects errors that have already occurred; a rationally organised workflow prevents a substantial proportion of them from occurring at all. Three points of intervention precede the post-editing stage (Table 5).

Table 5

Points of preventive intervention in the workflow

Point of intervention	**Measure**	**Expected effect**
Pre-editing of the source	Normalisation of script and of the letters o‘ and g‘; disambiguation of polysemous realia; simplification of overlong agglutinative chains	Elimination of orthographic errors and of part of the polysemy errors
Terminology management	A bilingual realia glossary with do-not-translate lists and forced terminology injected into the engine	Terminological consistency across the whole project
Prompt design (for generative models)	A concise direct instruction specifying the retention strategy and the register; supply of the surrounding context	Reduction of over-domestication and of hallucinated explication

The third point deserves particular comment. Comparative research indicates that concise, direct instructions outperform elaborate interpretive ones: an over-specified prompt induces the model to embellish, which in the case of realia manifests itself precisely as hallucinated explication. The practical corollary is that the instruction should state the strategy — retain the Uzbek term and supply a short gloss at first mention — rather than describe the desired stylistic effect at length.

The glossary occupies the central place in this system. Its function is not to relieve the post-editor of decision-making but to ensure that a decision taken once is reproduced consistently throughout the project and across projects. In this respect the terminology base performs for realia the role that the translation memory performs for repeated segments: it converts an individual solution into an institutional resource.

8. DISCUSSION

The material examined permits a general conclusion regarding the division of labour between the machine and the human being in this language pair. Contemporary systems cope satisfactorily with the syntactic and morphological transfer of Uzbek into English and, in the case of generative models, with much of the stylistic organisation as well. What they do not do is take responsibility for a cultural decision, because such a decision presupposes a model of the reader — an assessment of what the recipient knows, what needs explaining and what may be left implicit. This is exactly the competence that ISO 18587 ascribes to the post-editor.

It should also be noted that the improvement in machine quality does not diminish the requirements placed on the post-editor but redistributes them. When output was manifestly defective, post-editing was largely a matter of repair, and errors announced themselves. As fluency approaches the human level, the residual errors become semantic and cultural, that is to say invisible without the source text. Post-editing thereby ceases to be a simplified form of translation and becomes a form of expert evaluation, requiring higher rather than lower qualifications.

Finally, the low-resource status of Uzbek has a consequence that is easily overlooked. Poor corpus representation of realia means that machine output on such units is not merely of lower quality but less stable: the same unit may be rendered differently in different contexts within a single text. Consistency, therefore, cannot be expected of the engine and must be imposed externally, through terminology management. This is an argument in favour of institutional rather than individual solutions: a unified realia glossary for Uzbek cultural, tourist and literary content would remove a considerable share of the problems described here before any post-editing began.

9. CONCLUSIONS

1. Uzbek realia should be classified for machine-translation purposes not only by subject sphere but by degree of corpus representation; units with partial representation give rise to the most dangerous errors, since they are replaced by confident approximate analogues that give the reader no signal of loss.

2. Nine recurrent machine error types may be distinguished in the rendering of realia; they map onto the accuracy, terminology, linguistic-conventions and style categories of the MQM framework and divide into a highly detectable formal group and a poorly detectable semantic and cultural group.

3. The light post-editing regime defined in ISO 18587:2017 is structurally incapable of resolving realia-related errors, since its criterion is comprehensibility whereas the error lies in the cultural layer; projects intended to convey national specificity must be planned as full post-editing from the outset.

4. Realia load the cognitive component of post-editing effort above all; on culturally saturated material the correction of machine output may prove more labour-intensive than translation from scratch.

5. The repertoire of linguistic solutions comprises seven procedures whose selection is governed by the function of the unit in the text rather than by editorial preference; a three-step decision sequence — status, function, position — permits this choice to be formalised.

6. A significant proportion of errors can be prevented before post-editing through pre-editing of the source, terminology management with do-not-translate lists, and concise directive prompt design for generative models.

10. Practical recommendations

1. To compile a unified bilingual glossary of Uzbek realia with fixed English renderings, transliteration norms and usage notes, and to make it available for integration into machine translation and CAT environments.

2. To include in project specifications an explicit requirement of full post-editing for any content with a cultural component, with a corresponding revision of rates and deadlines.

3. To introduce a pre-editing stage normalising the script and the letters o‘ and g‘ and disambiguating polysemous realia in the source text.

4. To supplement automated quality assurance with a check for terminological consistency of realia units across the whole project, since this error class is formally detectable.

5. To include in the training of translators a module on post-editing that treats the recognition of low-detectability errors — over-domestication, hallucinated explication and polarity reversal — as a distinct skill.

References

1. ISO 18587:2017. Translation services — Post-editing of machine translation output — Requirements. – Geneva: International Organization for Standardization, 2017. – 19 p.

2. ISO 17100:2015. Translation services — Requirements for translation services. – Geneva: International Organization for Standardization, 2015. – 21 p.

3. MQM Council. The MQM error typology. – URL: https://themqm.org/error-types-2/typology/

4. Lommel A., Uszkoreit H., Burchardt A. Multidimensional Quality Metrics (MQM): A framework for declaring and describing translation quality metrics // Tradumàtica. – 2014. – №12. – P. 455–463.

5. Wang Q. Evaluating Uighur literary translation: A comparative study of ChatGPT, Google Translate, and Bing Translator // PLOS ONE. – 2025. – Vol. 20, №10. – Art. e0335261.

6. Krings H.P. Repairing texts: Empirical investigations of machine translation post-editing processes. – Kent: Kent State University Press, 2001. – 635 p.

7. O'Brien S. Towards a dynamic quality evaluation model for translation // The Journal of Specialised Translation. – 2012. – №17. – P. 55–77.

8. Koponen M. Is machine translation post-editing worth the effort? A survey of research into post-editing and effort // The Journal of Specialised Translation. – 2016. – №25. – P. 131–148.

9. Newmark P. A textbook of translation. – New York: Prentice Hall, 1988. – 292 p.

10. Vinay J.-P., Darbelnet J. Comparative stylistics of French and English: A methodology for translation. – Amsterdam: John Benjamins, 1995. – 358 p.

11. Baker M. In other words: A coursebook on translation. 3rd ed. – London: Routledge, 2018. – 390 p.

12. Pedersen J. Subtitling norms for television: An exploration focussing on extralinguistic cultural references. – Amsterdam: John Benjamins, 2011. – 242 p.

13. Venuti L. The translator's invisibility: A history of translation. 2nd ed. – London: Routledge, 2008. – 336 p.

14. Nida E.A. Toward a science of translating. – Leiden: E.J. Brill, 1964. – 331 p.

15. Влахов С., Флорин С. Непереводимое в переводе. – М.: Международные отношения, 1980. – 343 с.

16. Комиссаров В.Н. Современное переводоведение. – М.: ЭТС, 2002. – 424 с.

17. Ranathunga S., Lee E.-S.A., Skenduli M.P., et al. Neural machine translation for low-resource languages: A survey // ACM Computing Surveys. – 2023. – Vol. 55, №11. – P. 1–37.

18. Díaz Cintas J., Remael A. Subtitling: Concepts and practices. – London: Routledge, 2021. – 292 p.