



O'zbekiston Respublikasi
Maktabgacha va maktab
ta'limi vazirligi

ISSN: 2992-9008
UDK: 37

TA'LIM VA TARAQQIYOT

Ilmiy-uslubiy jurnali



2026-YIL 5-SON

www.journal.namspi.uz

Bosh muharrir: Namangan davlat pedagogika instituti rektori B.E.Xusanov
Mas'ul muharrir: Ilmiy ishlar va innovatsiyalar bo'yicha prorektor M.R.Qodirxonov
Mas'ul muharrir o'rinbosari: Ilmiy tadqiqotlar, innovatsiyalar va ilmiy-pedagogik kadrlar tayyorlash bo'limi boshlig'i J.A.Yuldashev

TAHRIR HAY'ATI

Aniq va tabiiy fanlar: k.f.d., prof. Sh.Abdullayev, f-m.f.d., dots. R.Xakimov, k.f.d., prof. N.Vohidova, k.f.d., dots. M.Qodirxonov, b.f.d., dots. A.Batashov, f-m.f.n., dots., T.Abdullayev, p.f.n., dots. M.Eshnazarova, p.f.n., dots. G'.Nafasov, p.f.n., dots. M.Boltayeva.

Texnika fanlari: t.f.d., prof. A.Umarov, t.f.d., prof. A.Qayumov, PhD, dots. Sh.Xudoyqulov, PhD, dots. Abdunazarov.

Ijtimoiy-gumanitar fanlar: s.f.d., prof. T.Fayzullayev, tar.f.d., prof. A.Rasulov. f.f.d., prof. M.Ismoilov, i.f.d., dots. N.Sotivoldiyev, t.f.n., dots. A.Isoqboyev, f.f.n., dots. O.Mamatov, t.f.n., dots. Z.Haydarov, PhD. J.Yuldashev, PhD. A.Abdullayev.

Pedagogika fanlari: p.f.d., prof. B.Xodjayev, p.f.d., prof. X.Najmiddinova, p.f.d., prof. T.Ismoilov, p.f.d., prof. S.Abdullayev. p.f.d., prof., M.Y.Sobirova, p.f.d., prof., R.Zamilova, p.f.d., prof. Sh.Xujamberdiyeva, p.f.d., dost. M.Asqarova, PhD, dots. M.Asranboyeva.

Filologiya fanlari: fil.f.d., prof. N.Uluqov, fil.f.d., prof. H.Usmanova, f.f.d., dots. X.Shokirova, p.f.n., dots. S.Misirov, p.f.n., dots. Sh.Ubaydullayev, PhD, dots. P.Lutfullayev, PhD, dots. M.Yakubbayev.

Texnik muharrir: R.Tursunov

Tahririyat manzili: Namangan shahri, Uychi ko'chasi, 316-uy.

Pochta indeksi: 160136, **Faks:** (0369) 227-27-21,

Email: Info@namspi.uz

"Ta'lim va taraqqiyot" ilmiy-uslubiy jurnali elektron to'plam sifatida O'zbekiston Respublikasi Prezidenti Administratsiyasi huzuridagi **Axborot va ommaviy kommunikatsiyalar agentligi (AOKA)ning 2022-yil 25-oktabrdagi № 045044 raqamli guvohnomasi** asosida chop etiladi. Jurnal **Oliy Attestatsiya komissiyasining 2024-yil 30-noyabrdagi (№ 364/5) yig'ilishi qaroriga asosan pedagogika fanlari bo'yicha falsafa fanlari doktori (PhD) va fan doktori (DSc) ilmiy darajasiga talabgorlarning dissertatsiya ishlari yuzasidan asosiy ilmiy natijalarini chop etish uchun tavsiya etilgan.** UDK: 37, Xalqaro standartlashtirish raqami (ISSN): 2992-9008.

Maqolalarning ilmiy saviyasi va keltirilgan ma'lumotlar uchun mualliflar javobgar hisoblanadi.

FROM CHOICE TO PROBABILITY: RETHINKING VENUTI'S DOMESTICATION IN THE MACHINE TRANSLATION OF UZBEK PROVERBS AND IDIOMS: AN EXPLORATORY PILOT STUDY

Bahriddinova Feruza Bahriddin qizi¹, Husainov Ilyos Jamoliddin o'g'li²

¹Department of Teaching English to Speakers of Other Languages (TESOL), Webster University
in Tashkent, Tashkent, Uzbekistan

E-mail: rbaxriddinova@webster.edu

²Department of Business Management, Turan International University, Namangan, Uzbekistan

E-mail: husainov_ilyos@turan-edu.uz

Abstract. This exploratory pilot study asks whether Lawrence Venuti's distinction between domesticating and foreignizing translation can describe how machine translation (MT) systems handle Uzbek proverbs and idioms. Venuti developed these concepts to describe the choices of human translators, mainly between European languages. The article first reviews his framework and describes the position of Uzbek as a "low-resource" language in natural language processing (NLP). It then compares how Google Translate and ChatGPT translated five Uzbek proverbs and idioms into English. The two systems produced similar results for three items: both replaced one Uzbek idiom with the established English idiom "kill two birds with one stone," and both translated two proverbs closely. They differed on the remaining two items, where ChatGPT substituted an English expression and Google Translate did not. These results suggest that the tendency to domesticate depends on two factors: how well established an English equivalent is, and the type of system. On this basis, the article proposes a distinction between strategic domestication, which is a translator's deliberate choice, and retrieval domestication, which occurs when a system reproduces an equivalent pair already present in the texts it learned from. It also argues that close, literal MT output is better described as default literalism than as foreignization, because it does not reflect any stance toward the foreign text. Given the small sample, these ideas are offered as hypotheses for larger studies.

Keywords: machine translation; large language models; domestication; foreignization; Uzbek; low-resource languages; proverbs; idioms

TANLOVDAN EHTIMOLLIKKA: O'ZBEK MAQOL VA IBORALARINI MASHINA ORQALI TARJIMA QILISHDA VENUTINING "MAHALLIYLASHTIRISH" (DOMESTICATION) TUSHUNCHASINI QAYTA KO'RIB CHIQISH – DASTLABKI TADQIQOT

Annotatsiya. Ushbu dastlabki tadqiqotda Lourens Venutining tarjimini "mahalliyashtirish" (domestication) va "chet elliklashtirish" (foreignization)ga ajratish yondashuvi mashina tarjimasini (MT) tizimlarining o'zbek maqol va iboralarini bilan ishlash jarayonini tavsiflashga qay darajada mos kelishi

o'rganiladi. Venuti ushbu tushunchalarni, asosan, Yevropa tillari o'rtasidagi tarjimalarda inson-tarjimonlarning tanlovlarini tasvirlash uchun ishlab chiqqan edi. Maqolada avvalo uning nazariy yondashuvi ko'rib chiqiladi hamda tabiiy tilni qayta ishlash (NLP) sohasida o'zbek tilining "kam resursli til" sifatidagi o'rni ta'riflanadi. So'ngra Google Translate va ChatGPT tizimlarining beshta o'zbek maqol va iborasini ingliz tiliga o'girishdagi natijalari taqqoslanadi. Ikkala tizim uchta holatda o'xshash natija ko'rsatdi: ular bitta o'zbek iborasini ingliz tilidagi "kill two birds with one stone" (bir o'q bilan ikki quyonni urish) iborasi bilan almashtirdi va ikkita maqolni o'xshash tarzda tarjima qildi. Qolgan ikki holatda esa farq kuzatildi: ChatGPT inglizcha iboradan foydalangan bo'lsa, Google Translate bunday qilmadi. Ushbu natijalar shuni ko'rsatadiki, "mahalliyashtirish"ga moyillik ikki omilga bog'liq: ingliz tilidagi ekvivalentning qay darajada keng tarqalganligi va tizim turi. Shu asosda maqolada "strategik mahalliyashtirish" (tarjimonning ongli tanlovi) hamda "qidiruv orqali mahalliyashtirish" (tizim o'zi o'rgangan matnlarda mavjud bo'lgan ekvivalent juftlikni takrorlashi natijasida yuzaga keladigan holat) o'rtasidagi farq taklif etiladi. Shuningdek, mashina tarjimasining so'zma-so'z (literal) natijasi "chet elliklashtirish" emas, balki "standart so'zma-so'zlik" (default literalism) deb atalishi to'g'riroq ekanligi ta'kidlanadi, chunki u asl matnga nisbatan hech qanday maxsus yondashuv yoki pozitsiyani aks ettirmaydi. Tanlanma hajmi kichik bo'lgani sababli, ushbu fikrlar kengroq ko'lamli tadqiqotlar uchun gipoteza sifatida ilgari suriladi.

Kalit so'zlar: mashina tarjimasi; katta til modellari; mahalliyashtirish; chet elliklashtirish; o'zbek tili; kam resursli tillar; maqollar; iboralar

ОТ ВЫБОРА К ВЕРОЯТНОСТИ: ПЕРЕОСМЫСЛЕНИЕ КОНЦЕПЦИИ ДОМЕСТИКАЦИИ ЛОУРЕНСА ВЕНУТИ В КОНТЕКСТЕ МАШИННОГО ПЕРЕВОДА УЗБЕКСКИХ ПОСЛОВИЦ И ИДИОМ (ПИЛОТНОЕ ИССЛЕДОВАНИЕ)

Аннотация. В данном пилотном исследовании рассматривается вопрос о том, применимо ли разграничение Лоуренса Венути между доместикацией (одомашниванием) и фореинизацией (очуждением) к тому, как системы машинного перевода (МП) обрабатывают узбекские пословицы и идиомы. Венути разработал эти концепции для описания стратегий выбора, к которым прибегают переводчики-люди, работая преимущественно с европейскими языками. В статье сначала анализируется теоретический подход Венути и описывается статус узбекского языка как языка с ограниченными ресурсами (low-resource language) в сфере обработки естественного языка (NLP). Затем сравниваются варианты перевода пяти узбекских пословиц и идиом на английский язык, выполненные системами Google Translate и ChatGPT. Для трех выражений системы выдали схожие результаты: обе заменили одну узбекскую идиому на устойчивый английский аналог «kill two birds with one stone» (убить двух зайцев одним выстрелом) и выполнили близкий к оригиналу перевод двух пословиц. Различия проявились в двух оставшихся случаях: ChatGPT использовал английское выражение-эквивалент, тогда как Google Translate этого

не сделал. Полученные результаты позволяют предположить, что склонность к доместикации зависит от двух факторов: степени устоявшеести английского эквивалента и типа используемой системы. На этом основании в статье предлагается различать «стратегическую доместикацию» (осознанный выбор переводчика) и «доместикацию на основе извлечения» (воспроизведение системой пары эквивалентов, уже имеющейся в обучающих текстах). Также утверждается, что буквальный перевод, выдаваемый системой, правильнее называть «буквализмом по умолчанию», а не фореинизацией, поскольку он не отражает какой-либо позиции по отношению к иноязычному тексту. Учитывая малый объем выборки, данные идеи выдвигаются в качестве гипотез для более масштабных исследований.

Ключевые слова: машинный перевод; большие языковые модели; доместикация; фореинизация; узбекский язык; языки с ограниченными ресурсами; пословицы; идиомы

INTRODUCTION

How far a translation should adapt a foreign text to the reader's culture is one of the oldest questions in translation studies. The distinction goes back to Friedrich Schleiermacher's (1813/2012) argument that a translator may move the reader toward the author or the author toward the reader, and it is usually discussed today through Lawrence Venuti's (1995) related distinction between domesticating and foreignizing translation. A domesticating translation brings the foreign text close to the target culture and reads fluently, so that readers do not notice they are reading a translation; a foreignizing translation keeps the text's difference visible, even at the cost of fluency. For Venuti, this is not only a technical matter but an ethical one: a fluent, domesticating translation adapts the foreign text to the values of the receiving culture, and in the Anglo-American context this contributes to a wider cultural dominance in which foreign texts are accepted only once they stop sounding foreign. Both strategies, in his account, are ethical stances a translator takes toward the foreign. The framework, however, assumes a translator who makes choices. Machine translation (MT) challenges this assumption directly: MT systems and chatbots are now commonly used to translate culturally embedded language, turning foreign images into familiar ones or leaving them unchanged, and it is not clear whether such outputs can be called domestication or foreignization in Venuti's sense when no one has made a deliberate choice.

Venuti's framework has been influential but also widely criticized. Critics have questioned how consistently the domestication–foreignization binary can be applied to actual translation decisions, and have argued more broadly that the framework's assumptions do not transfer easily outside the Anglo-American literary contexts in which it was developed; others have noted that foreignization can shade into exoticism rather than genuine respect for difference. This article does not try to resolve these debates. It uses domestication and foreignization because they remain the standard terms for describing how a translation relates to the foreignness of its source, even though they were developed with human translators and mostly European language pairs in mind. Research on MT usually evaluates systems in terms of accuracy and adequacy, often with automatic metrics, and work on MT for Turkic languages, including Uzbek, has focused mainly on building systems and measuring their quality (e.g., Mirzakhlov et al., 2021). Idioms are a known problem in this research: Dankers et al. (2022) show that neural MT systems often fail to translate idioms correctly and tend to produce overly literal, word-by-word translations, though their study examined translation from English into European languages rather than from a low-resource language into English. More recent work suggests that system type may matter for how literal an output is. Raunak et al. (2023) ask directly whether GPT-based models produce less literal translations than conventional neural MT systems, a question that anticipates the system-type contrast this article's results later raise. Castaldo and Monti (2024) likewise show that prompting strategy shapes how large language models render idiomatic expressions, which bears on how the present study's own prompts should be reported and interpreted (see 2.4). Research that connects MT output directly with Venuti's concepts is less common. Jones and Irvine (2013) analyzed statistical MT translations of French literature from the perspective of translation theory and argued that MT systems cannot be said to choose between domestication and foreignization, because they reproduce patterns found in their training data; they also observed that idioms which appear in the training data are

easy for a system to reproduce. This is the core problem the present article takes up: Venuti's framework presupposes a translator who takes a stance toward the foreign, and an MT system, which has no stance, cannot straightforwardly be described in these terms. This question is especially relevant for languages such as Uzbek. In NLP, Uzbek is described as a low-resource language, meaning that far less digital text and parallel data are available for it than for English. In a survey of the Turkic languages of Central Asia, Veitsman and Hartmann (2025) rank Uzbek second after Kazakh in data availability and technology development, and, using the resource categories proposed by Joshi et al. (2020), suggest that Uzbek could be placed among "the Hopefuls" if data collection continues. They trace the region's data scarcity to three causes: the continuing influence of Russian in science, education, and public life since the Soviet period; limited internet access; and the absence of long-term investment in language technology. Uzbek's history of script change — from an Arabic-based script to Latin, then Cyrillic, and back to Latin after independence, with both scripts still in use today — means the available textual record is itself a product of this history rather than a neutral technical resource. Studies of Uzbek–English proverb translation by human translators offer a further point of comparison (Kadyrova, 2020; Niyadullayeva, 2025).

The present article builds on these observations in three ways: it examines current systems, including a chatbot based on a large language model, rather than statistical MT; it focuses on a low-resource language rather than French; and it proposes specific terms for the different kinds of output such systems produce, addressing a gap in terminology for choice-less MT output. A reasonable starting expectation was that systems trained on much more English than Uzbek would push Uzbek expressions toward English norms in broadly similar ways. Proverbs and idioms are a useful test case for this expectation, because they carry cultural images that a translation can either keep or replace. This article accordingly reports an exploratory pilot study of five Uzbek proverbs and idioms translated into English by Google Translate and ChatGPT, and asks two questions. RQ1

(empirical): How do the two systems translate these items, and do they tend toward domestication or close translation? RQ2 (conceptual): Can Venuti's concepts describe these outputs, and if not, how should they be adjusted? RQ1 is addressed in the Results; RQ2 is addressed in the Discussion. The article does not claim to establish general patterns; its aim is to propose a conceptual distinction that larger studies can test.

METHODS

This is an exploratory pilot study with a qualitative design. A five-item qualitative comparison is appropriate at this stage because the study is hypothesis-generating and aims at concept development rather than frequency estimation; its purpose is not to measure how often domestication occurs but to examine individual outputs closely enough to propose concepts for further research.

Five items were selected by the authors, who are native speakers of Uzbek: three proverbs (maqol) and two idioms (ibora). Item selection was guided by the authors' informal judgment of a range of imagery and apparent closeness to existing English expressions; formal equivalence status was coded only after translations were collected (see 2.5, 2.6). Source items were entered in Latin script. Wording was verified against O'zbek xalq maqollari (2-jildlik; A. Madvaliyev et al.) and O'zbek tilining iboralar lug'ati (Sh. Rahmatullayev).

Two source words carry a polysemy relevant to interpreting the outputs: o'q, which in contemporary Uzbek most commonly means "bullet" but historically, and in related Turkic languages such as Turkish (ok), means "arrow"; and urmoq, which in hunting contexts can mean "to shoot" as well as "to hit."

Two systems were used. The first was Google Translate, web version, a dedicated translation service, with the source language set manually to Uzbek. The second was ChatGPT 5.1, Pro tier, a general-purpose chatbot based on a large language model.

The two systems should not be treated as a simple contrast between traditional neural MT and a large language model. In December 2025, Google announced that Translate

would use its Gemini models to improve translation of idioms, slang, and other expressions with nuanced meaning, first in the United States and India for English and nearly 20 other languages (Yao, 2025); whether this change applied to Uzbek at the time of the present study is not known. For this reason, the study compares a dedicated translation service with a general-purpose chatbot without making claims about their internal architecture.

Each item was entered once into each system in August 2026, without surrounding context. Google Translate received the bare Uzbek string. ChatGPT received the following prompt, with the Uzbek idiom or proverb inserted in place of the placeholder:

You are an expert linguist specializing in Uzbek lexicography, phraseology, and translation studies.

Your task is to translate and explain the following Uzbek idiom(s) or proverb(s) into English for an academic and cultural comparison study.

INPUT DATA

Uzbek Idiom/Proverb: "[INSERT_UZBEK_IDIOM_OR_PROVERB_HERE]"

INSTRUCTIONS

Please provide a structured output following these four distinct sections:

- 1. Literal Translation: Provide a word-for-word translation of the phrase into natural English, preserving the direct syntax and imagery.*
- 2. Figurative/Idiomatic Meaning: Explain the actual meaning, tone, and cultural context of the phrase in simple, clear English. Mention when and how it is typically used in Uzbek culture (e.g., formal/informal, warning, advice, sarcasm).*
- 3. English Equivalent(s): Provide 1 to 3 standard English idioms, proverbs, or expressions that carry the exact or closest equivalent meaning, tone, and functional usage.*
- 4. Linguistic & Cultural Nuance Analysis: Identify any key cultural metaphors, historical roots, or semantic nuances unique to the Uzbek language in this phrase that do not translate directly into English.*

EVALUATION CRITERIA FOR YOUR RESPONSE

Your output will be benchmarked on: accuracy of the literal vs. figurative mapping; relevance and naturalness of the English equivalents; depth of cultural insight regarding Uzbek heritage and usage context; structural precision and consistency with the requested layout.

Only the literal translation and, where relevant, the English-equivalent output of this structured response are reported in Table 2 and analyzed in the Results and Discussion; the figurative-meaning and cultural-nuance sections of ChatGPT's output are not separately analyzed in this study.

Each output was classified using a coding scheme adapted from Baker's (2018) strategies for translating idioms, with one additional category — close translation — introduced by the authors to capture outputs with no domesticating substitution. Substitution is adapted from Baker's strategy of replacing an idiom with a target-language idiom of similar meaning but different form. Paraphrase follows Baker's strategy of expressing the meaning without an idiom. Hybrid and close translation are the authors' own additions: hybrid combines a paraphrase or idiom with a literal gloss of the source image, and close translation, which may partly correspond to Baker's strategy of borrowing the source-language idiom, keeps the image and structure of the source. In this article, domestication means adapting a source text to the norms and cultural references of the target language so that its foreignness is reduced; foreignization means keeping features of the source text even when this makes the translation less fluent. Substitution is the clearest form of domestication. The mapping of close translation onto Venuti's foreignizing pole is formal, however, and is revisited in the Discussion (4.3).

Category	Definition	Decision rule	Formal position on domestication–foreignization continuum
Substitution	Source expression replaced by a target-language idiom of similar meaning, different form	Output uses an established English idiom/proverb in place of the source image	Domesticating
Paraphrase	Meaning expressed without an idiom	Output states the meaning in ordinary language; no idiom present	Domesticating
Hybrid	Paraphrase or idiom combined with a literal gloss of the source image	Output contains both a meaning-based rendering and the source image	Mixed
Close translation	Source image and structure retained	Output preserves the source's central image and syntactic structure	Formally foreignizing

Table 1. Coding scheme.

Two further dimensions were coded for each item: (a) equivalence status (established, partial, none), based on listing in named Uzbek–English or Russian–English phraseological dictionaries; (b) presence of a Russian parallel conventionally paired with the same English expression; (c) whether the source image was retained; and (d) whether meaning was fully or partially retained. Naturalness was not coded systematically and is

not reported as a finding in the Results; where it is mentioned in the Discussion, it reflects the authors' native or near-native judgment.

All classifications were made by a single annotator (the first author), and the coding was independently reviewed by the second author. Equivalence status, Russian-parallel status, and the other secondary dimensions in 2.5 were coded after the translations had been collected.

RESULTS

Table 2 presents the five items, their literal glosses, and the outputs of both systems. Table 3 shows how each output was classified, together with equivalence status and the presence of a Russian parallel.

No.	Uzbek item (type) and literal gloss	ChatGPT	Google Translate
1	Aytilgan so'z – otilgan o'q (proverb) "A spoken word – a shot arrow/bullet"	"A word once spoken is like an arrow once shot"	"A spoken word is like a fired arrow"
2	Boshi osmonga yetmoq (idiom) "For one's head to reach the sky"	"To be over the moon"	"To be overjoyed (to have one's head reach the sky)"
3	Bir o'q bilan ikki quyonni urish (idiom) "To hit two hares with one shot"	"To kill two birds with one stone"	"To kill two birds with one stone"

No.	Uzbek item (type) and literal gloss	ChatGPT	Google Translate
4	Har qush uyasida ko'rganini qiladi (proverb) "Every bird does what it has seen in its nest"	"A bird does what it sees in its nest"	"Every bird does in its nest what it has seen done there"
5	Do'st boshga kulfat tushganda sinaladi (proverb) "A friend is tested when misfortune falls on one's head"	"A friend in need is a friend indeed"	"A friend is tested when trouble strikes"

Table 2. Uzbek items and machine translation outputs.

No.	ChatGPT	Google Translate	Agree?	Equivalence status	Russian parallel?	Image retained?
1	Close translation	Close translation	Yes	None	No	Yes (both)
2	Substitution	Hybrid	No	Partial	No	No (ChatGPT); yes (Google, in gloss)
3	Substitution	Substitution	Yes	Established	Yes	No (both)
4	Close translation	Close translation	Yes	Partial	No	Yes (both)

No.	ChatGPT	Google Translate	Agree?	Equivalence status	Russian parallel?	Image retained?
5	Substitution	Close translation	No	Established	Yes	No (ChatGPT); yes (Google)

Table 3. Classification of the outputs, with equivalence status and Russian-parallel and image-retention coding.

The systems agreed on three items (1, 3, and 4) and differed on two (2 and 5). ChatGPT used substitution for three of the five items, while Google Translate used it only once, for item 3. In both cases where the systems differed, ChatGPT's output was the more domesticating one.

Item 1: both systems rendered o'q as "arrow" and preserved the central image and structure of the source; both were coded close translation. The only change was the addition of "like," which turns the Uzbek metaphor into a simile.

Item 3: both systems produced the identical output "to kill two birds with one stone," coded substitution for both. The source image (hares, a single shot) was fully replaced by the target image (birds, a stone). This item is coded as having an established equivalent with a Russian parallel (Table 3).

Item 4: both systems produced close translations. Neither used "the apple doesn't fall far from the tree" or the closer proverb "as the old cock crows, so crows the young." This item is coded as having a partial equivalent. The two outputs differ in surface fluency: the Google Translate output ("every bird does in its nest what it has seen done there") has less conventional English syntax than the ChatGPT output.

Item 2: ChatGPT substituted "to be over the moon" (coded substitution); Google Translate produced "to be overjoyed (to have one's head reach the sky)" (coded hybrid), giving the meaning first and the literal image in parentheses. This item is coded as a partial equivalent with no Russian parallel recorded.

Item 5: ChatGPT substituted “a friend in need is a friend indeed” (coded substitution); Google Translate produced “a friend is tested when trouble strikes” (coded close translation). The Uzbek images of testing and of misfortune “falling on the head” are not retained in the ChatGPT output. This item is coded as having an established equivalent with a Russian parallel (Table 3).

DISCUSSION

The starting expectation was that both systems would domesticate Uzbek expressions in broadly similar ways. The results only partly support this. The systems agreed completely when a strong, established equivalent was available (item 3) and also agreed when no clear equivalent was available, translating closely (items 1 and 4). They differed on items 2 and 5, and in both cases ChatGPT domesticated while Google Translate did not. RQ1 can therefore be answered as follows: both systems produced domesticating output under some conditions, but not uniformly, and the degree of domestication depended on both the item and the system.

This pattern suggests two factors. The first is how well established an equivalent pair is: when a pair such as “two hares with one shot” / “two birds with one stone” is common in bilingual texts, both systems appear to reproduce it.

The second factor is system type. ChatGPT substituted an English expression whenever a reasonably close one existed, while Google Translate did so only when the equivalent was very strong. This is broadly consistent with Raunak et al.'s (2023) finding that GPT-based models can produce less literal output than conventional neural MT, and with Dankers et al.'s (2022) observation that literalness is a common default in neural MT more generally; it also fits Jones and Irvine's (2013) account of MT systems reproducing patterns dominant in their training data, extended here from statistical MT to a general-purpose chatbot. One plausible explanation is that a general-purpose chatbot is designed to produce natural, fluent English, while a dedicated translation service stays closer to the source text — that is, the two systems may have different thresholds for substitution.

An alternative, non-exclusive explanation is the prompting difference between the two systems. As reported in 2.4, ChatGPT was given a detailed, structured prompt that explicitly asked for standard English equivalents and cultural-nuance analysis, while Google Translate received only the bare Uzbek string; this contextual framing, rather than architecture alone, could account for part of ChatGPT's greater tendency to substitute. Castaldo and Monti (2024) show that prompting strategy measurably affects how large language models render idiomatic expressions, so this confound cannot be ruled out on the basis of the present design, and disentangling it from system type is a priority for future research (4.6).

Item 4 also bears on this explanation. The closer English proverb, “as the old cock crows, so crows the young,” is rare in modern usage, which may explain why neither system produced it even though it is a closer semantic match than “the apple doesn't fall far from the tree.” The hybrid output for item 2 (Google Translate) is unusual for a dedicated translation service and resembles the explanatory style associated with large language models; it may be connected to the Gemini update described in 2.3, but this cannot be confirmed with the available evidence.

In response to RQ2: for Venuti, domestication is a strategic and ethically significant choice — a human translator chooses fluency and, with it, the removal of the foreign. The outputs examined here are not strategic in this sense. Neither system takes a position toward the foreign text; each produces the output that its training data makes most probable. This is consistent with Jones and Irvine's (2013) argument that MT systems cannot handle domestication and foreignization in Venuti's sense.

The results therefore suggest a distinction between two kinds of domestication. Strategic domestication is the deliberate choice Venuti describes. Retrieval domestication is a domesticating output that appears when a system reproduces an equivalent pair already established in the texts it learned from. “Retrieval” is used here in its general sense of recalling a stored pattern; it does not refer to retrieval-augmented generation, a

specific technique in NLP. Retrieval domestication clearly explains item 3. It may also explain ChatGPT's outputs for items 2 and 5, if that system treats less-established pairs as equivalents more readily than Google Translate does.

The same logic applies to the other pole of Venuti's binary. When a system translates an item closely, as in items 1 and 4, the output may look foreignizing, especially when its syntax is unnatural. However, foreignization in Venuti's sense is also a deliberate stance — a decision to keep the foreign visible. A literal output produced because no equivalent is available is not such a decision. This article therefore proposes calling it default literalism: the term describes what the output does without implying an intention behind it, and is consistent with Dankers et al.'s (2022) finding that neural MT systems tend to translate idioms too literally.

Venuti's main objection to domestication is that it makes translation invisible by hiding the foreign origin of a text behind a fluent surface. Once this effect is named, however, readers can recognize it and ask what choices the translator made. Retrieval domestication is invisible in a different and more complete way: there is no translator whose choices can be questioned, and the output gives no sign that a substitution has taken place.

In items 3 and 5, the substituted English proverbs keep most of the meaning of the Uzbek originals, but the cultural images disappear — the hunter's shot, the hares, and the misfortune that “falls on the head.” For item 5 specifically, two things are lost in the ChatGPT substitution beyond the image: the target proverb, “a friend in need is a friend indeed,” is grammatically ambiguous, since “a friend in need” can also be read as a friend who needs help, whereas the Uzbek proverb carries no such ambiguity. An English reader has no way of knowing that the Uzbek expression differs and may reasonably conclude that Uzbek simply has the same proverb as English. The risk is greater when two proverbs are only partly equivalent, as item 4 illustrates: had a system produced “the apple doesn't fall far from the tree,” the meaning would have shifted from learned

behavior to inherited character, and nothing in the output would reveal the change. Retrieval domestication has no built-in check for such differences.

Google Translate's hybrid output for item 2 suggests a possible middle path: by giving both the English meaning and the literal Uzbek image, it makes the act of translation visible and restores some of the visibility that fluent translation removes, even though no translator made that decision. Whether this format appears consistently, and whether readers find it useful, are questions for future research.

This study has several limitations. The sample was small (five items) and was selected by the authors rather than drawn systematically from a proverb collection, so the findings cannot be assumed to generalize. Each item was translated only once; MT outputs can vary between queries and change over time, so the results describe a single moment and may not be reproducible. Classifications were made by a single annotator and independently reviewed by the second author. Equivalence status and the other secondary coding dimensions were judged after the translations were collected rather than fixed in advance; this creates a risk of circular reasoning, in which substitution is taken as evidence that an equivalent pair exists, and the pair is then used to explain the substitution. No published human translations were used as a baseline. The exact technology behind Google Translate at the time of the study — specifically, whether the Gemini update described in 2.3 applied to Uzbek — is uncertain. Finally, as noted in 2.4 and 4.2, ChatGPT received a detailed structured prompt while Google Translate received only the bare source string; this asymmetry means that prompt wording, rather than system architecture alone, may account for some of the observed differences between the two systems.

The findings suggest three testable hypotheses for larger studies. First, substitution is more likely for items with dictionary-attested English equivalents than for items without them. Second, general-purpose chatbots have a lower substitution threshold than dedicated translation services, producing domesticating output for a

wider range of source items. Third, Uzbek items with Russian–English counterparts are substituted more often than items without them, reflecting the historical influence of Russian on the region's bilingual textual record. A larger study of 20–30 items, classified for equivalence status in advance using bilingual dictionaries, could test these hypotheses. Each item could be translated several times and by several systems, including Yandex Translate, which is widely used in the region alongside Google Translate (Veitsman & Hartmann, 2025). A second annotator could classify outputs independently, and published human translations could provide a baseline.

CONCLUSION

This article asked whether Venuti's distinction between domestication and foreignization can describe how MT systems translate Uzbek proverbs and idioms into English. The pilot study suggests the concepts are useful but need adjustment. Both systems produced domesticating outputs, but not uniformly: domestication appeared most clearly when a strong English equivalent was available, and the general-purpose chatbot substituted English expressions more readily than the translation service. To describe these outputs without assuming a translator's intention, the article proposed two terms: retrieval domestication, as distinct from strategic domestication, and default literalism, as distinct from foreignization. These terms are offered as hypotheses rather than findings, given the study's small sample and other limitations. Connecting the distribution of training data to the linguistic history that produced it — including the influence of Russian and Uzbek's changes of script — may be one of the most productive directions for this line of research.

References

1. Baker, M. (2018). *In other words: A coursebook on translation* (3rd ed.). Routledge.
2. Castaldo, A., & Monti, J. (2024). Prompting large language models for idiomatic translation. In *Proceedings of the 1st Workshop on Creative-text Translation and Technology* (pp. 32–39). European Association for Machine Translation. <https://aclanthology.org/2024.ctt-1.4/>
3. Dankers, V., Lucas, C., & Titov, I. (2022). Can Transformer be too compositional? Analysing idiom processing in neural machine translation. In *Proceedings of the 60th Annual Meeting of*

- the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 3608–3626). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2022.acl-long.252>
4. Jones, R., & Irvine, A. (2013). The (un)faithful machine translator. In Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (pp. 96–101). Association for Computational Linguistics. <https://aclanthology.org/W13-2713>
5. Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 6282–6293). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.560>
6. Kadyrova, M. (2020). Uzbek proverbs and their English equivalents: A cultural-linguistic study. Tashkent University Press.
7. Mirzakhlov, J., Babu, A., Ataman, D., Kariev, S., Tyers, F., Abduraufov, O., Hajili, M., Ivanova, S., Khaytbaev, A., Laverghetta, A., Jr., Moydinboyev, B., Onal, E., Pulatova, S., Wahab, A., Firat, O., & Chellappan, S. (2021). A large-scale study of machine translation in Turkic languages. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (pp. 5876–5890). Association for Computational Linguistics.
8. Niyadullayeva, S. (2025). Translating idioms and proverbs across cultures. World Scientific Research Journal, 40(1), 348–353.