HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

# REDEFINING AFFECTIVE FIDELITY: A FRAMEWORK FOR MEASURING EMOTIONAL EQUIVALENCE IN THE TRANSLATION OF ENGLISH SENTIMENTAL POETRY INTO UZBEK

**Akhatov Maksud Sharofiddin ugli**

Master Student, Study Program Linguistik (S2), Universitas Pendidikan Indonesia

Email: maksudakhatov@upi.edu

Supervisor: **Iwa Lukmana**

Email: iwa@upi.edu

Doi: https://doi.org/10.5281/zenodo.24202725

**Abstract:** Studies of translation quality have long asked whether a translation keeps the meaning of its source and have seldom asked whether it keeps the feeling. That gap matters more now that neural and large-language-model (LLM) systems produce poetry translations that are fluent and broadly accurate, while research on literary machine translation continues to report losses in voice, creativity, and emotion. This article reviews earlier definitions of affective or emotional fidelity in simulation research and machine translation evaluation, argues that they are too coarse for poetry, and redefines *affective fidelity* as the degree to which a translation reproduces, through the resources of the target language, the kind, intensity, and texture of the emotion encoded in a source poem. The concept draws on equivalence theory, poetry translation theory, appraisal theory, and dimensional and categorical models of emotion, and the article applies it to English sentimental poetry translated into Uzbek. Six dimensions make up the resulting framework: lexical-affective, morphological, prosodic, imagistic-cultural, pragmatic-tonal, and reader-response. The article gives particular attention to the grammatical features through which Uzbek carries emotion, such as diminutive and endearment suffixes, possessive intimacy, emotive particles, and syllabic (*barmoq*) metre. The framework is operationalised as a scoring rubric, the Affective Fidelity Index (AFI), together with a typology of affective loss, and contrastive renderings of lines by Thomas Gray and Robert Burns illustrate its use. No empirical scores are reported. Instead, the article supplies the instrument that a comparative study of human and AI translations would need and states how that instrument should be validated.

**Keywords:** affective fidelity; poetry translation; emotional equivalence; machine translation; large language models; sentimentalism; Uzbek; translation quality assessment

## 1. Introduction

A love poem can be translated accurately and still fail. Suppose Burns's speaker compares his beloved to a rose, and the translation renders the comparison correctly but in the tone of a botanical note. The reader then learns what the poem says without feeling what it does. That second property, the emotional effect, is what readers of poetry most often come for, and it is also the property that translation assessment is least equipped to measure.

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

The question is now a practical one. Native speakers sometimes judge literary translations produced by neural machine translation (NMT) and large language models (LLMs) to be equal to professional work at sentence level (Toral & Way, 2018), and LLMs that translate whole paragraphs make fewer errors than systems working sentence by sentence (Karpinska & Iyyer, 2023). The same studies, and others, nonetheless report that critical errors remain and that a human is still needed to keep the author's voice intact (Karpinska & Iyyer, 2023). In reader-response research, human translations of fiction score higher than machine-translated and post-edited versions for creativity, narrative engagement, and reception (Guerberof-Arenas & Toral, 2020), and post-editing constrains the translator's creativity (Guerberof-Arenas & Toral, 2022). Work on poetry in particular includes lack of emotion among the problems of LLM output (Wang et al., 2024).

Existing metrics have no name for the loss these findings describe. Automatic measures such as BLEU score the overlap between a translation and a reference (Papineni et al., 2002), and neural metrics such as COMET predict human quality judgements (Rei et al., 2020), yet neither was designed to tell whether a translated elegy still grieves. Human evaluation schemes usually cover adequacy and fluency, or creativity and engagement, without asking which emotion a poem carries and whether that emotion survives. Translation studies has begun to take emotion seriously, though mostly as a property of translators and their decision-making (Hubscher-Davidson, 2018; Lehr, 2021) and not as an assessable property of the translated text. Recent computational studies have started to measure emotion in translated text (Zhou et al., 2026; Wiśniewski & Czudy, 2026). As Section 2.5 argues, however, they define it too narrowly to capture most of what carries feeling in a poem.

For language pairs with few resources, the gap is wider. Uzbek has a long tradition of literary translation and a school of translation theory of its own (Salomov, 1983), but computational work on the language is recent and is still building basic datasets (Mamasaidov et al., 2025). Much of the emotional meaning in Uzbek is also carried by morphology, in diminutive and endearment suffixes, possessive forms, and particles that an English source does not mark in the same way (Bodrogligeti, 2003). A system that maps English words onto Uzbek words may therefore produce a correct sentence and still omit exactly the forms that an Uzbek reader hears as tenderness or sorrow.

The article responds to this problem with a concept and an instrument. It reviews earlier definitions, redefines *affective fidelity*, builds a six-dimension framework for assessing it, and turns that framework into a scoring rubric with an error typology. English sentimental poetry serves as the test case because the genre exists to produce feeling, and Uzbek serves as the target language because it realises feeling through means that English lacks. The article addresses four research questions:

RQ1. How have earlier studies defined affective or emotional fidelity, and how should it be redefined so that it is distinct from semantic and formal equivalence and suited to poetry?

RQ2. Through which linguistic levels does English sentimental poetry carry affect, and which corresponding resources does Uzbek offer?

RQ3. How can affective fidelity be operationalised as a replicable assessment instrument?

RQ4. How does the instrument distinguish emotionally strong from emotionally weak renderings, as shown on illustrative lines?

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

The study is conceptual, and it reports no scores from real translations or readers. Section 5 illustrates the framework on contrastive renderings constructed for the purpose, and Section 7 sets out how the instrument should be validated empirically.

## 2. Literature Review

### *2.1 Equivalence and the reader's response*

The idea that a translation should reproduce an effect as well as a meaning can be traced back at least to Nida (1964). His dynamic equivalence required the target reader to respond to the translation substantially as the source reader had responded to the original. Nida placed the reader at the centre without defining the response itself, and he neither separated understanding from feeling nor said how either should be measured. Jakobson (1959) held that poetry is untranslatable in the strict sense, because in poetry the sound and grammatical form of words carry meaning, which leaves only creative transposition possible. If, as Nida held, the response matters, and if, as Jakobson held, form carries part of it, then the emotional part of the response needs an account of its own. This article takes up that problem.

### *2.2 Poetry translation theory*

For the most part, poetry translation theory has classified strategies, not effects. Comparing English translations of Catullus, Lefevere (1975) set out seven strategies, among them phonemic, literal, metrical, and rhymed translation, and showed that each one gives up some dimension of the poem in order to keep another. For Holmes (1988), a verse translation is a *metapoem*, a new poem that interprets the old one, and the formal choices translators make shape how the poem is received. Both accounts imply that a translation may keep a poem's content and lose its effect. Neither, however, offers a way of assessing the emotional effect on its own terms.

### *2.3 Emotion in language and in translation*

Two families of emotion models can be borrowed from psychology. Categorical models name a small set of basic emotions. Plutchik (2001) proposes eight of them (joy, trust, fear, surprise, sadness, disgust, anger, anticipation), which vary in intensity and combine into more complex states. Dimensional models instead place feelings in a continuous space, and Russell's (1980) circumplex arranges affective states around two axes, pleasure–displeasure (valence) and activation (arousal). Within a translation framework, the categorical view identifies *which* emotion a line carries, while the dimensional view indicates *how strongly* and *in which direction*.

Linguistics contributes an account of how language encodes feeling. Appraisal theory, which developed within Systemic Functional Linguistics, divides evaluative language into attitude (affect, judgement, appreciation), engagement, and graduation, and graduation describes how meanings are intensified or softened (Martin & White, 2005). For the present purpose graduation is particularly useful, since a translation can keep the type of an emotion while weakening or exaggerating its force.

When translation studies turned to emotion in recent years, it did so chiefly from the translator's side. Hubscher-Davidson (2018) identifies three places where emotions enter translation, namely the emotional material of the source text, the translator's own emotions, and the emotions of source and target readers. Experimental work shows that induced emotions affect translation performance and

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

creativity (Rojo López & Ramos Caro, 2016), and reviews have mapped how emotion and cognition interact in the translation process (Lehr, 2021). Hubscher-Davidson's first and third places, the emotion in the text and the emotion in its readers, remain less developed, and they are the subject of this article.

### *2.4 Machine and LLM translation of literary texts*

The question asked of literary machine translation has shifted from whether it works to what it loses. Toral and Way (2018) trained NMT on more than 100 million words of literary text and found that, depending on the novel, 17–34% of its sentences were judged equivalent to professional translation. According to Guerberof-Arenas and Toral (2020, 2022), professional translation done without MT scored higher for creativity, engagement, and reception, and post-editing constrained creativity. Karpinska and Iyyer (2023) showed that LLMs make fewer errors when translating whole paragraphs than sentence-level systems do, although critical mistranslations and omissions persist.

Poetry presents the hardest case. Chakrabarty et al. (2021) found that general MT systems often fail to preserve the meaning and poetic style of verse and that fine-tuning on poetic data helps. Wang et al. (2024) identified persistent problems in ChatGPT's English–Chinese poetry translations and improved them by supplying explanations of the poem as guidance, and their review of earlier work lists lack of emotion among the weaknesses of LLMs. After comparing LLM and human poetry translations for lexical variety, density, and rhyme, Resende and Hadley (2024) concluded that the effort of post-editing may outweigh the benefits of using the machine at all. Yet none of these studies measures emotion as a separate construct, and evaluation continues to rely on BLEU, COMET, or general human ratings (Papineni et al., 2002; Rei et al., 2020).

### *2.5 Earlier definitions of affective fidelity*

Neither *affective fidelity* nor its near-synonym *emotional fidelity* is a new term. The concept has been defined in at least three ways, which Table 1 summarises. Each definition suits the purpose it was made for, but each is too narrow for poetry translation.

The first definition originates outside translation studies. Writing on simulated training environments, Harris et al. (2020) place affective fidelity alongside physical and psychological fidelity and state that it "requires the simulation to elicit a realistic emotional response in the user". On this view, fidelity is a property of the person experiencing the simulation and not of the simulation itself, and it is measured through self-reports and physiological signals such as heart rate and skin conductance. In spirit the definition is close to Nida's (1964) dynamic equivalence, and it has the same weakness. It indicates whether the response is right but not which features of the artefact produced it.

The second sense is computational. Zhou et al. (2026) describe emotional fidelity as "a distinct, quantifiable, and reference-free dimension of MT quality assessment". Their measure, EMOS, compares source and target emotion vectors over Ekman's six basic emotions plus a neutral category, and it combines vector similarity, a match of the dominant emotion label, and a ratio of emotional diversity with weights of 0.50, 0.35 and 0.15. When EMOS was applied to Classical Chinese–English translations by Google Translate, DeepL and GPT-4o, the dominant emotion was kept in only about 71–74% of segments. In one case, a translation that BLEU (0.734) and BERTScore (0.825) rated highly received a low EMOS score (0.427). Wiśniewski and Czudy (2026) follow a similar route

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

under the name *emotion preservation*. They translate English Reddit comments into five European languages and back, run emotion classifiers trained on 22 GoEmotions categories on the results, and treat the drop in macro-F1 as "affective bleaching". On average the drop was 2.9–4.9 percentage points. In the worst configuration desire, fear and anger lost most, and prompts that asked the model to preserve emotion made almost no difference.

In the third sense, loss is treated as drift. Al-Araji and Hasan (2026) define emotional drift as "the diversion or transformation of emotional signals from source text to its translations" and propose detecting it in the Arabic and English versions of the *Epic of Gilgamesh* through sentiment and emotion analysis.

**Table 1. Earlier definitions of affective or emotional fidelity**

| Source | Field | Where fidelity is located | How it is measured | Gap for poetry translated into Uzbek |
|---|---|---|---|---|
| Harris et al. (2020) | Simulation and training | The user's emotional response | Self-report; heart rate, skin conductance, EEG | No account of the features that cause the response |
| Zhou et al. (2026), EMOS | MT evaluation | Distribution of emotion labels in source and target | Vector similarity, label match and diversity over Ekman's six emotions plus neutral | Coarse categories; assumes equivalent emotion spaces across languages |
| Wiśniewski & Czudy (2026) | MT evaluation | Emotion that survives a round trip back to English | Drop in classifier macro-F1 after back-translation (22 GoEmotions categories) | Scores the English back-translation, not the target text its readers meet |
| Al-Araji & Hasan (2026) | Literary NLP | Shift in emotional signals between versions | Sentiment and emotion analysis | Signals at the level of sentiment; no treatment of form |

When the text is a poem and the target language is Uzbek, these definitions share four problems. The first is that their categories are too coarse for literary feeling, since tender and passionate love both count as joy, and quiet melancholy and anguished grief both count as sadness. The second is that they read emotion from words. A classifier registers *sevgilim* ('my beloved') but cannot tell whether tenderness is carried by a suffix such as *-jon*, by rhythm, or by an image, and these are precisely the places where Uzbek can carry it (Bodrogligeti, 2003). The third is that the computational methods assume reliable emotion models in both languages. Zhou et al. (2026) acknowledge that the cross-lingual equivalence of their emotion spaces is an assumption, and Wiśniewski and Czudy (2026) sidestep the question by scoring back-translations in English, so neither study measures what a target-language reader receives. For Uzbek, whose basic translation resources are still being built (Mamasaidov et al., 2025), that assumption is harder to justify. The fourth problem is that readers and texts are kept apart. Harris et al. (2020) measure the experiencer alone, and the computational studies measure the text alone, apart from the human ratings that EMOS used to set its weights.

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

These accounts nonetheless establish points worth keeping. EMOS distinguishes the kind of an emotion from its strength. The back-translation study shows that loss is uneven across emotions and that simply asking a model to keep the emotion achieves little, and Harris et al. (2020) insist that fidelity is in the end a matter of what the experiencer feels. The redefinition in Section 4 retains these points and adds what poetry requires, namely texture, the grammatical and prosodic means of the target language, and an explicit link between textual features and readers' responses.

### *2.6 The Uzbek context*

Literary translation in Uzbek has a strong tradition and a theory of its own. Salomov (1983) laid the foundations of Uzbek translation theory, and leading poets translated world classics, the best-known example being Erkin Vohidov's Uzbek *Faust* (Kun.uz, 2016). Within this tradition a poetic translation is a creative act, judged by its effect on Uzbek readers, and this makes the tradition a natural benchmark for machine output. Burns has a place in it as well. Muhammad Ali made the first Uzbek translations of his poems, and Xolbekova, Sultonova and Abdusamatova published further translations after independence (Davletova, 2024).

Two features of Uzbek make emotion especially hard to carry across in translation. Uzbek is agglutinative, and it encodes evaluative and affective meanings in suffixes and particles, among them diminutive and endearment forms and possessive markers of intimacy (Bodrogligeti, 2003), whereas English often expresses the same feelings through word choice, through intonation, or not at all. Uzbek is also still a low-resource language for NLP, and its core translation datasets have been built only recently (Mamasaidov et al., 2025), so systems have seen comparatively little Uzbek poetry. Taken together, the two features lead one to predict that machine translations into Uzbek will tend to render English emotion lexically, if at all, and will make too little use of the morphological resources a human poet would reach for. No published study has yet tested this prediction, and testing it requires an instrument of the kind proposed below.

## 3. Method

### *3.1 Research design*

This is a conceptual study that builds a framework, and it proceeds in three steps. The first step reviews earlier definitions critically (Section 2.5) and redefines affective fidelity so that it is distinct from semantic and formal equivalence (RQ1). The second maps the linguistic levels at which English sentimental poetry carries affect against the resources Uzbek offers at each level, and derives the dimensions of the framework from that mapping (RQ2). In the third step, each dimension is turned into scoring criteria and an error typology, which yields a rubric that trained raters can apply and whose reliability can be checked (RQ3).

### *3.2 Illustrative material*

Section 5 shows how the instrument works (RQ4) by applying it to three lines from two canonical sentimental poems in the public domain, Thomas Gray's *Elegy Written in a Country Churchyard* (1751) and Robert Burns's *A Red, Red Rose* (1794). Gray's *Elegy* represents the melancholic, meditative strand of sentimentalism, and Burns's lyric its tender, song-like strand. Janet Todd (1986)

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

describes eighteenth-century sensibility as privileging spontaneous emotion, tenderness, and sympathy, and the two poems display these qualities in compact form.

Two Uzbek renderings of each line are compared. Rendering A is a close lexical rendering, which keeps the propositional content and uses no affective morphology beyond what the English words require. Rendering B is oriented towards affect and draws on Uzbek affective resources. The author constructed both for this article in order to isolate the features the framework assesses. Neither is the output of a particular AI system or a published human translation, and they are offered to demonstrate the rubric, not as evidence of how any system or translator behaves.

### *3.3 Limits of the design*

A conceptual article cannot establish that the six dimensions are exhaustive, that the rubric is reliable, or that rater scores predict reader responses. Those are empirical questions, and Section 7 proposes a design for answering them. The study can, however, state the construct clearly enough for that validation to be designed.

## 4. The Affective Fidelity Framework

### *4.1 A redefinition*

As Section 2.5 showed, earlier definitions place affective fidelity either in the experiencer's response (Harris et al., 2020) or in a distribution of emotion labels over the text (Zhou et al., 2026; Wiśniewski & Czudy, 2026). The redefinition given below keeps both places and adds the linguistic means that link them.

*Affective fidelity* is the degree to which a translation reproduces the emotion encoded in a source text, in its kind, its intensity, and its texture, through means that are natural in the target language and culture. The concept lies close to three existing ones without being identical to any of them. Semantic equivalence has no way of registering that a translation may state the same proposition with a different emotion, or with none, whereas affective fidelity registers exactly that. In contrast to formal equivalence, affective fidelity does not require the same devices, so an English emotion carried by intonation may be faithfully rendered by an Uzbek suffix. Its nearest relative is Nida's (1964) dynamic equivalence. Nida, however, treats the reader's response as a whole, while affective fidelity isolates the emotional part of that response and makes it measurable.

Each of the three components can be stated precisely. *Kind* is the emotion category, defined through Plutchik's (2001) basic emotions and their blends; resignation, for example, is made up of sadness and acceptance. *Intensity* is the force of the emotion, which appraisal theory treats as graduation (Martin & White, 2005) and the circumplex treats as arousal (Russell, 1980). *Texture* is the particular colouring that an emotion takes in a given text, as in tender versus passionate love or quiet versus anguished grief, and it depends on register, imagery, and rhythm. Kind corresponds roughly to the dominant-label match in EMOS and intensity to its vector similarity (Zhou et al., 2026). Texture has no counterpart in the earlier definitions.

### *4.2 Six dimensions*

Affective fidelity is assessed at six levels. Dimensions 1–5 are text-internal, and trained raters score them by comparing source and translation. Dimension 6 is text-external and is measured by asking

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

target readers how the translation makes them feel. Because the two kinds of evidence are kept separate, the analyst can check whether what the text does matches what readers report.

**Table 2. Dimensions of affective fidelity**

| Dimension | What it assesses | Typical English carriers | Uzbek resources to check |
|---|---|---|---|
| 1. Lexical-affective | Whether emotion-laden words keep their valence and intensity | Evaluative adjectives, emotion nouns, archaic or elevated diction | Choice among neutral and marked synonyms, including Turkic and Persian-Arabic layers (e.g., *qabr* vs. *go'r*, 'grave') |
| 2. Morphological | Emotion carried by word structure | Few markers; occasional diminutives (*-ie*, *-let*) | Diminutive and endearment suffixes (*-cha*, *-gina*, *-jon*), possessives of intimacy (*-im*), intensive reduplication (*qip-qizil*) |
| 3. Prosodic | Rhythm, metre, and sound patterns that create mood | Accentual-syllabic metre, rhyme, repetition, alliteration | Syllabic *barmoq* metre, rhyme, assonance, refrain |
| 4. Imagistic-cultural | Whether images evoke equivalent feelings for target readers | Pastoral, graveyard, and floral imagery | Symbolic systems of Uzbek poetry (garden, rose, nightingale, separation) |
| 5. Pragmatic-tonal | Speaker's attitude and stance toward subject and addressee | Apostrophe, exclamation, interjection, address forms | Emotive particles (*-ku*, *-da*, *axir*), vocatives, address forms, exclamatory syntax |
| 6. Reader-response | Emotional effect on target readers | Measured empirically | Reader ratings of valence, arousal, and emotion category |

English and Uzbek differ most in Dimensions 1 and 2. Where English often carries tenderness lexically or prosodically and leaves the grammar neutral, Uzbek can carry it in a suffix that has no English counterpart (Bodrogligeti, 2003). A translation that renders only what the English words mark will therefore be grammatically correct and yet affectively thinner than an Uzbek poem would normally be.

### *4.3 The Affective Fidelity Index (AFI)*

For each unit of analysis, whether a line or a stanza, Dimensions 1–5 are each scored on a five-point scale. Raters judge the translation against the source on all three components of emotion, namely kind, intensity, and texture.

**Table 3. Scoring scale for Dimensions 1–5**

| Score | Label | Descriptor |
|---|---|---|

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

| Score | Label | Descriptor |
|---|---|---|
| 4 | Full | Kind, intensity, and texture reproduced through natural target-language means |
| 3 | Strong | Kind and intensity reproduced; texture slightly altered |
| 2 | Partial | Kind reproduced; intensity noticeably weakened or exaggerated |
| 1 | Weak | Emotion barely perceptible, or shifted to a neighbouring emotion |
| 0 | Lost | Emotion absent or replaced by a different emotion |

Scores for Dimension 6 come from reader data. Following the two axes of Russell's (1980) circumplex, readers rate each unit for valence and arousal on nine-point scales, and they choose the emotion they perceive from a list based on Plutchik (2001). Readers of English, or bilingual readers, give the same ratings for the source text, so that the distance between source and translation can be computed.

The index for a unit combines the dimension scores:

$$AFI = [\ \Sigma\ (w_i \times s_i) \div (4 \times \Sigma\ w_i)\ ] \times 100, \text{ where } i = 1 \ldots 6$$

In this formula, $s_1$ to $s_5$ are the rater scores (0–4), $s_6$ is the reader-response score rescaled to 0–4, and $w_i$ are the dimension weights. The weights should be equal ($w_i = 1$) in a first study, and unequal weights should be introduced only if empirical work shows that some dimensions predict reader response better than others. AFI ranges from 0 to 100. Where reader data are not available, a text-internal index (AFI-T) can be computed from Dimensions 1–5 alone and reported under that name.

Raters should be trained on the rubric with practice items and should score independently, without knowing which translation is human and which is machine. Their agreement should be reported with a reliability coefficient such as Krippendorff's alpha.

### *4.4 A typology of affective loss*

Whereas scores show how much emotion is kept, a typology shows how it is lost. Five types of affective shift are proposed, and each can be recorded together with its dimension. A study could then report, for example, that machine translations show mostly morphological flattening while human translations show occasional intensification.

**Table 4. Types of affective shift**

| Type | Definition | Example of the pattern |
|---|---|---|
| Flattening | Same emotion, lower intensity | Tender address rendered as neutral reference |
| Shifting | Emotion changes category | Resignation rendered as despair |
| Intensification | Same emotion, higher intensity than the source | Quiet melancholy rendered as lament |
| Tonal mismatch | Emotion kept, register or stance wrong | Intimate address rendered in formal register |
| Cultural | Image kept literally but evokes nothing, or | A source image with no resonance in |

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

| Type | Definition | Example of the pattern |
|---|---|---|
| opacity | something else, for target readers | the target tradition left unadapted |

Intensification is included because affective fidelity does not mean maximal emotion. A translator who adds sentiment that the source lacks has also departed from it, and the rubric should penalise that departure in the same way as it penalises loss.

## 5. Illustrative Application

The framework is applied here to three lines. In each case, Rendering A keeps the propositional content with neutral grammar, and Rendering B draws on Uzbek affective resources. As Section 3.2 explains, the author constructed both renderings for demonstration. The scores in Table 5 are therefore the author's worked examples of the rubric and not research data, and Dimension 6 is not scored because no readers were consulted.

### *5.1 Burns: tender comparison*

> O my Luve's like a red, red rose, That's newly sprung in June; (Burns, *A Red, Red Rose*, 1794)

Rendering A: *Oh, mening sevgim qizil, qizil atirgulga o'xshaydi, / Iyunda yangi ochilgan.*

Rendering B: *Yorim — iyunda yangi ochilgan / Qip-qizil gul misoli, jonim.*

The source emotion is tender, joyful love: positive valence, moderate arousal, with a folk-song texture created by the Scots spelling and the exclamatory *O*. At the lexical level (Dimension 1), Rendering A's *sevgim* ('my love') can mean the feeling as well as the person, while *yor* in Rendering B is the conventional word for the beloved in Uzbek love lyric. At the morphological level (Dimension 2), the English repetition *red, red* is an intensifier; Rendering A copies it as *qizil, qizil*, which is unusual in Uzbek, while Rendering B uses the native intensive reduplication *qip-qizil*. The vocative *jonim* ('my soul, my dear') in Rendering B compensates for the exclamation *O*, which Rendering A keeps as a borrowed interjection (Dimension 5). The rater must still ask whether *jonim* adds tenderness the source lacks, a possible case of intensification.

### *5.2 Burns: farewell*

> And fare thee weel, my only Luve! (Burns, *A Red, Red Rose*, 1794)

Rendering A: *Va xayr, mening yagona sevgim!*

Rendering B: *Alvido, yagonam, jonginam!*

The source emotion is a blend of love and sorrow at parting. Rendering A's *xayr* is the everyday word for goodbye, which creates a tonal mismatch with a solemn farewell; *alvido* in Rendering B is the elevated, poetic farewell. Morphologically, Rendering B places the possessive on the adjective (*yagonam*, 'my only one') and doubles the endearment in *jon-gina-m*, carrying in suffixes what Burns carries in the intimate *thee* and in dialect. Here too the rater must judge texture: *-gina* can lend a diminutive sweetness that may tilt Burns's manly folk tenderness toward something softer.

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

### *5.3 Gray: melancholy resignation*

> The paths of glory lead but to the grave. (Gray, *Elegy Written in a Country Churchyard*, 1751)

Rendering A: *Shon-shuhrat yo'llari faqat qabrga olib boradi.*

Rendering B: *Shon-shuhrat yo'li go'rga eltar, axir.*

The source emotion is resigned melancholy: negative valence, low arousal, and a gnomic, timeless texture. Rendering A states the proposition in neutral prose, with the standard *qabr* for 'grave' and the ordinary present *olib boradi*. Rendering B uses *go'r*, a sombre Turkic word for the grave, and the aorist *eltar*, which gives the statement the timeless quality of a proverb (Dimension 1). The particle *axir* ('after all') adds a stance of weary acceptance that matches Gray's *but* (Dimension 5), though in some contexts it can sound argumentative. Prosodically (Dimension 3), Rendering B has eleven syllables, a common *barmoq* line length, while Rendering A runs to fifteen syllables of unmetred prose.

### *5.4 Worked scoring*

**Table 5. Illustrative AFI-T scores (author's worked examples; Dimension 6 not measured)**

| Line | A/B | D1 | D2 | D3 | D4 | D5 | AFI-T | Shift types noted |
|---|---|---|---|---|---|---|---|---|
| Burns, rose | A | 2 | 1 | 1 | 2 | 1 | 35 | Flattening (D2, D5) |
| Burns, rose | B | 4 | 4 | 3 | 3 | 3 | 85 | Possible intensification (D5) |
| Burns, farewell | A | 2 | 1 | 1 | 2 | 2 | 40 | Tonal mismatch (D1); flattening (D2) |
| Burns, farewell | B | 4 | 3 | 3 | 3 | 3 | 80 | Possible texture shift (D2) |
| Gray, glory | A | 2 | 2 | 1 | 3 | 1 | 45 | Flattening (D3, D5) |
| Gray, glory | B | 3 | 3 | 3 | 3 | 3 | 75 | Possible tonal mismatch (D5) |

D1 = lexical-affective; D2 = morphological; D3 = prosodic; D4 = imagistic-cultural; D5 = pragmatic-tonal. AFI-T = sum of D1–D5 ÷ 20 × 100.

Three features of the instrument emerge from the worked examples. It separates renderings that a semantic adequacy check would treat as equal, even though all six keep the propositional content. Its dimension profile also locates the difference, showing that Rendering A loses most in Dimensions 2, 3, and 5, which are precisely where Uzbek carries emotion that English leaves to other means. Finally, the shift typology keeps the rubric from rewarding emotion as such, since the stronger renderings attract flags of their own for possible intensification or change of texture. Whether real AI translations into Uzbek resemble Rendering A, and real human translations resemble Rendering B, is an empirical question that the instrument is designed to answer and that this illustration does not settle.

## 6. Discussion

The article's central claim is that emotion in translation can be treated as a construct with components, levels, and failure modes of its own, and not as an impression left over once accuracy has been checked. Defining affective fidelity through kind, intensity, and texture answers RQ1 by giving raters three questions that semantic equivalence does not ask. The mapping of the six dimensions answers

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

RQ2 and shows that a poem does not carry emotion at the same levels in English as in Uzbek. RQ3 is answered by translating the dimensions into a five-point rubric, an index, and a shift typology, and the worked examples in Section 5 address RQ4 by showing that the instrument separates renderings that a meaning-based check would treat as equivalent.

The framework links two bodies of research that have so far developed in parallel. Studies of literary MT have repeatedly observed losses of voice, creativity, and emotion (Guerberof-Arenas & Toral, 2020, 2022; Karpinska & Iyyer, 2023; Wang et al., 2024), yet they have measured these losses with general scales of quality, engagement, or creativity. Research on translation and emotion has built a psychological account of emotion in translation (Hubscher-Davidson, 2018; Lehr, 2021), but its focus has been on translators. Affective fidelity borrows the emotion models of the second line of research and applies them to the textual product that the first line evaluates.

The framework yields a specific, testable prediction for English–Uzbek translation. Since Uzbek carries much of its affect in morphology and particles (Bodrogligeti, 2003), and since systems have seen relatively little Uzbek poetry (Mamasaidov et al., 2025), machine translations should lose most in Dimensions 2 and 5, with flattening as their dominant shift type. Human translators trained in the Uzbek tradition (Salomov, 1983) should score higher on those dimensions, although they may show more intensification, since an Uzbek poetic register tends toward rich endearment. Should the prediction hold, the dimension scores will show where the gap between human and machine translation lies, and so tell developers and post-editors what to fix.

Beyond research, the framework suggests practical uses. Post-editors of machine output could treat Dimensions 2 and 5 as a checklist for emotional repair. Teachers of literary translation could have students score their own drafts on the rubric, which makes explicit a judgement of tone that usually remains tacit. Prompting research on LLMs, such as the explanation-guided method of Wang et al. (2024), could also include affective instructions and measure their effect with the AFI instead of BLEU.

Three risks remain. Emotional judgement is subjective, and raters from different backgrounds may disagree, particularly about texture; training and reliability checks reduce this risk without removing it. Because the framework privileges the target reader's experience, it may conflict with a foreignizing strategy in which strangeness is intended. Finally, the boundary between legitimate compensation and intensification is not sharp. The *jonim* in Section 5.1 can be read either way, and the rubric records the question instead of resolving it.

**7. Conclusion**

This article has argued for affective fidelity as a separate dimension of translation quality and has built an instrument for assessing it. Affective fidelity is defined as the reproduction, through natural target-language means, of an emotion's kind, intensity, and texture, and it is assessed at six levels: lexical-affective, morphological, prosodic, imagistic-cultural, pragmatic-tonal, and reader-response. The Affective Fidelity Index combines rater and reader scores into a single figure, while a typology of five shift types (flattening, shifting, intensification, tonal mismatch, and cultural opacity) records how emotion is lost or distorted. When applied to English sentimental poetry and Uzbek, the framework

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

draws attention to the morphological and pragmatic resources through which Uzbek carries feeling, resources that a word-for-word rendering is likely to miss.

The study has a number of limitations. Because it is conceptual and its illustrative renderings were constructed by the author, it shows how the instrument works but not how real human or AI translations perform. The six dimensions derive from theory, and they may need to be merged, split, or supplemented once data have been collected. For want of evidence, the weights in the index are set equal. The illustrations, finally, cover three lines from two poets and cannot represent the range of English sentimental poetry.

Validation is the next step. A first study should (a) select a corpus of complete English sentimental poems, such as Gray's *Elegy* and several of Burns's lyrics, together with any published Uzbek translations and the outputs of two or three AI systems, recording system names, versions, dates, and prompts; (b) train at least three bilingual raters on the rubric and report inter-rater reliability with Krippendorff's alpha (Krippendorff, 2019); (c) collect blind ratings of valence, arousal, and emotion category from native Uzbek readers and, for the source, from English or bilingual readers; and (d) test whether the text-internal scores (AFI-T) predict the reader-response scores. If they do, AFI-T can serve as a practical proxy where reader studies are not feasible. If they do not, the mismatch will reveal which dimensions need revision. Either outcome would give translators and developers a measure of the emotional loss that readers already report in machine-translated poetry.

## References

1. Al-Araji, Z., & Hasan, B. T. (2026). Lost in translation? Detecting emotional drift in bilingual texts of Gilgamesh using NLP. *AlKadhim Journal for Computer Science, 4*(1), 12–21. https://doi.org/10.61710/kjcs.v4i1.146
2. Bodrogligeti, A. J. E. (2003). *An academic reference grammar of modern literary Uzbek* (Vols. 1–2). LINCOM Europa.
3. Chakrabarty, T., Saakyan, A., & Muresan, S. (2021). Don't go far off: An empirical study on neural poetry translation. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (pp. 7253–7265). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.emnlp-main.577
4. Davletova, Sh. (2024). Robert Byorns ijod yo'li [The journey of Robert Burns as a poet]. *Philology Matters, 50*(3), 143–162. https://doi.org/10.36078/987655018
5. Guerberof-Arenas, A., & Toral, A. (2020). The impact of post-editing and machine translation on creativity and reading experience. *Translation Spaces, 9*(2), 255–282. https://doi.org/10.1075/ts.20035.gue
6. Guerberof-Arenas, A., & Toral, A. (2022). Creativity in translation: Machine translation as a constraint for literary texts. *Translation Spaces*. https://doi.org/10.1075/ts.21025.gue
7. Harris, D. J., Bird, J. M., Smart, P. A., Wilson, M. R., & Vine, S. J. (2020). A framework for the testing and validation of simulated environments in experimentation and training. *Frontiers in Psychology, 11*, Article 605. https://doi.org/10.3389/fpsyg.2020.00605
8. Holmes, J. S. (1988). *Translated! Papers on literary translation and translation studies*. Rodopi.
9. Hubscher-Davidson, S. (2018). *Translation and emotion: A psychological perspective*. Routledge.

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

10. Jakobson, R. (1959). On linguistic aspects of translation. In R. A. Brower (Ed.), *On translation* (pp. 232–239). Harvard University Press. https://doi.org/10.4159/harvard.9780674731615.c18
11. Karpinska, M., & Iyyer, M. (2023). Large language models effectively leverage document-level context for literary translation, but critical errors persist. In *Proceedings of the Eighth Conference on Machine Translation* (pp. 419–451). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.wmt-1.41
12. Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). SAGE.
13. Kun.uz. (2016, May 31). *Erkin Vohidov olamdan o'tdi* [Erkin Vohidov has passed away]. https://kun.uz/news/2016/05/31/erkin-voidov-olamdan-utdi
14. Lefevere, A. (1975). *Translating poetry: Seven strategies and a blueprint*. Van Gorcum.
15. Lehr, C. (2021). Translation, emotion and cognition. In F. Alves & A. L. Jakobsen (Eds.), *The Routledge handbook of translation and cognition* (pp. 294–309). Routledge. https://doi.org/10.4324/9781315178127-20
16. Mamasaidov, M., Aral, A., Shopulatov, A., & Inomjonov, M. (2025). Filling the gap for Uzbek: Creating translation resources for Southern Uzbek. In *Proceedings of the Tenth Conference on Machine Translation* (pp. 1081–1087). Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.wmt-1.83
17. Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan. https://doi.org/10.1057/9780230511910
18. Nida, E. A. (1964). *Toward a science of translating, with special reference to principles and procedures involved in Bible translating*. Brill.
19. Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). BLEU: A method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics* (pp. 311–318). https://doi.org/10.3115/1073083.1073135
20. Plutchik, R. (2001). The nature of emotions. *American Scientist, 89*(4), 344–350.
21. Rei, R., Stewart, C., Farinha, A. C., & Lavie, A. (2020). COMET: A neural framework for MT evaluation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. https://doi.org/10.48550/arXiv.2009.09025
22. Resende, N., & Hadley, J. (2024). The translator's canvas: Using LLMs to enhance poetry translation. In *Proceedings of the 16th Conference of the Association for Machine Translation in the Americas (Volume 1: Research Track)* (pp. 178–189). AMTA. https://aclanthology.org/2024.amta-research.16
23. Rojo López, A. M., & Ramos Caro, M. (2016). Can emotion stir translation skill? Defining the impact of positive and negative emotions on translation performance. In R. Muñoz Martín (Ed.), *Reembedding translation process research* (pp. 107–130). John Benjamins. https://doi.org/10.1075/btl.128.06roj
24. Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178. https://doi.org/10.1037/h0077714
25. Salomov, G'. (1983). *Tarjima nazariyasi asoslari* [Foundations of translation theory]. O'qituvchi.

HSR (LONDON),
HOUGHTON STREET REVIEW JOURNAL
ISSN: 2058-5519|Impact factor-15.9
http://houghtonstreetreview.com/index.php/hsr

26. Todd, J. (1986). *Sensibility: An introduction*. Methuen.
27. Toral, A., & Way, A. (2018). What level of quality can neural machine translation attain on literary text? In J. Moorkens, S. Castilho, F. Gaspari, & S. Doherty (Eds.), *Translation quality assessment: From principles to practice* (pp. 263–287). Springer. https://doi.org/10.1007/978-3-319-91241-7_12
28. Wang, S., Wong, D. F., Yao, J., & Chao, L. S. (2024). What is the best way for ChatGPT to translate poetry? In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 14025–14043). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.acl-long.756
29. Wiśniewski, D., & Czudy, I. (2026). *Beyond semantics: Measuring fine-grained emotion preservation in small language model-based machine translation* [Preprint]. arXiv. https://arxiv.org/abs/2604.27920
30. Zhou, J., Cheng, F., & Zhang, X. (2026). EMOS: A multidimensional emotional fidelity evaluation framework for machine translation quality assessment. *Cognitive Computation, 18*(1), Article 65. https://doi.org/10.1007/s12559-026-10609-3
31. Burns, R. (1794). *A red, red rose*. Public domain.
32. Gray, T. (1751). *Elegy written in a country churchyard*. Public domain.