

Automating intercultural communication: pragmatic failures and cultural bias in AI-mediated Indonesian–English translation

Heratul Aini
Universitas PGRI Sumenep, Indonesia

Article Info

Article history:

Received Apr 26, 2026
Revised May 19, 2026
Accepted Jun 30, 2026

Keywords:

artificial intelligence;
cultural bias;
intercultural communication;
machine translation;
pragmatic failure;

ABSTRACT

Background: AI-mediated translation increasingly facilitates Indonesian–English communication, yet fluent outputs may obscure pragmatic failure and culturally reductive representation in institutionally produced public discourse across tourism, education, public information, and everyday digital interaction in Indonesia. **Objective:** This study examines how automated translation preserves interpersonal meaning, represents culturally embedded Indonesian expressions, and maintains consistency across systems and repeated generations. **Method:** A comparative qualitative corpus design analysed twelve Indonesian source segments from verified public institutional documents using a Pragmatic Equivalence Matrix, Cultural Representation Audit, and Cross-System Stability Test across forty-eight matched outputs. **Results:** Findings showed that interactional stance permeated all source segments, while audience alignment, mitigated directives, and figurative meanings created recurrent pressure points for automated translation. Cultural vulnerability centred on *mudik*, *gotong royong*, culinary naming, gift practices, and ritual symbolism, with generalisation and Anglocentric reframing compressing religious, communal, regional, and symbolic associations. **Implication:** Cross-system comparison indicated 56.3% strategy agreement, 60.4% lexical or framing variation, 37.5% recurrent pragmatic shifts, 31.3% repeated cultural reframing, and 70.8% severity consistency, demonstrating that stable surface quality did not ensure intercultural adequacy. **Novelty:** This study contributes an integrated evaluative framework that reconceptualises AI translation quality as the combined preservation of pragmatic force, cultural accountability, and cross-run stability rather than lexical accuracy and fluency alone.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Heratul Aini
Universitas PGRI Sumenep, Indonesia
Jl. Trunojoyo, Gedung Barat, Gedung, Kec. Batuan, Sumenep, Jawa Timur, 69451, Indonesia
Email: heranuraini123@gmail.com

1. INTRODUCTION

Artificial intelligence increasingly mediates multilingual encounters in Indonesia, where digital communication reaches a population of exceptional linguistic and cultural diversity. APJII reported 221,563,479 internet users in 2024, equivalent to 79.5% of Indonesia's population, creating extensive conditions for automated translation across tourism, education, commerce, public information, and interpersonal communication. Yet greater accessibility does not guarantee intercultural adequacy. Indonesian expressions such as *mudik*, *halal bihalal*, *gotong royong*, *buah tangan*, kinship address, religious formulae, and discourse particles encode social relations that cannot always be recovered through lexical substitution.

Reviews of AI-based translation likewise identify persistent difficulties involving context, culturally specific meaning, and low-resource languages, even as fluency improves [1], [2]. When fluent English obscures altered politeness, weakened illocutionary force, or reductive framing, users may trust outputs whose communicative consequences remain inaccurate. Automated Indonesian–English translation therefore constitutes a consequential site where cultural meaning is selected, normalized, or erased.

Recent scholarship has established foundations for examining this problem. Cross-cultural NLP research demonstrates that multilingual coverage does not ensure cultural representativeness, because models inherit uneven datasets, culturally bounded categories, and dominant-language assumptions [2], [3]. Studies of conversational AI foreground intercultural adaptation, trust, bias, and communication training [4], [5], [6], [7]. Translation-focused work documents cultural interference in Indonesian–English transfer and strategies for rendering culturally nuanced Indonesian texts [8], [9], while IndoCulture reveals geographically differentiated cultural commonsense across eleven Indonesian provinces [10]. However, these strands remain fragmented. Existing studies rarely combine segment-level pragmatic analysis, cultural-representation auditing, and repeated cross-system comparison using a verifiable Indonesian public-text corpus [11], [12]. Consequently, little is known about whether fluent AI translations preserve interpersonal meaning consistently or reproduce bias through omission, generalization, substitution, and Anglocentric reframing.

Addressing this gap, this study investigates how AI-mediated Indonesian–English translation manages pragmatically consequential and culturally embedded expressions drawn from a documented corpus of Indonesian institutional texts. It asks three questions. First, how do AI systems preserve or recalibrate speech acts, politeness, address forms, deixis, implicature, discourse particles, figurative language, stance, and interpersonal register when translating Indonesian segments into English? Second, which translation strategies are applied to culture-specific concepts, and when do retention, glossing, paraphrase, functional equivalence, generalization, substitution, literalization, or omission produce culturally reductive representations? Third, how stable are these decisions across systems, model versions, prompts, and repeated runs? To answer them, this study aligns three analytical instruments: a Pragmatic Equivalence Matrix, a Cultural Representation Audit, and a Cross-System Stability Test. Rather than treating grammatical accuracy or fluency as proxies for quality, this design evaluates translation as situated intercultural action dependent on communicative force, social positioning, and culturally anchored meaning.

This study advances the provisional argument that AI-mediated translation is vulnerable not where Indonesian expressions are formally incomprehensible, but where culturally dense meanings become fluent, plausible, and socially altered English. Such alteration may operate through three mechanisms. Pragmatic recalibration may neutralize address, soften or intensify directives, remove discourse particles, and explicate implicatures in ways that reorganize interpersonal relations. Cultural reframing may render local concepts intelligible by replacing them with broader English categories, yet simultaneously flatten regional, religious, historical, or symbolic specificity. Cross-run variability may show that these choices are not stable properties of a system output but probabilistic decisions sensitive to prompts and model configurations. This proposition extends critiques of cultural bias in language models [6], [13] and intercultural-pragmatic caution regarding AI-assisted learning [14]. Its implication is that trustworthy translation evaluation must examine cultural accountability and pragmatic equivalence alongside accuracy, fluency, and consistency.

2. LITERATURE REVIEW

2.1. Pragmatic failure

Pragmatic failure denotes a mismatch between linguistic form and intended meaning, producing communication that may be grammatically acceptable yet socially inappropriate. Foundational pragmatics distinguishes pragmlinguistic failure, arising from unsuitable form–function mappings, from sociopragmatic failure, arising from misjudged norms, relationships, or situational expectations. Recent intercultural research broadens this distinction by treating failure as relational rather than exclusively speaker deficient, because interpretation emerges through cultural negotiation [15], [16], [17]. In AI-mediated translation, this relational view competes with accuracy-centred models that privilege semantic correspondence. Accordingly, pragmatic adequacy requires preserving communicative force, interpersonal positioning, and culturally recognizable contextual assumptions.

Operationally, pragmatic failure can be examined through recurrent features that organize Indonesian interaction. These include speech acts, politeness orientation, address and kinship forms, deixis, implicature, presupposition, discourse particles, figurative expressions, stance, and register. Alterations may appear as intensified directives, weakened refusals, deleted honorifics, overexplicit implicatures, literalized idioms, or shifts from affiliative to impersonal voice [18], [19], [20]. Studies of culturally nuanced Indonesian translation further show that lexical transfer alone cannot preserve locally embedded interactional values [8], [9]. Together, these indicators therefore support a segment-level assessment of pragmatic equivalence systematically.

2.2. Cultural bias

Cultural bias in language technology refers to representational asymmetries whereby cultural assumptions become normalized while others are simplified, marginalized, or misrecognized. Some scholarship locates bias primarily in datasets, annotation practices, and model architecture; other work emphasizes bias as an interactional effect emerging when users interpret culturally framed outputs [2], [6], [13]. Cultural awareness is therefore not synonymous with factual knowledge about traditions. It also concerns whether systems recognize cultural values, relational categories, symbols, and expectations [3], [10]. This distinction shifts analysis from cultural mention toward representational consequence.

Cultural bias becomes observable through translation choices that reshape specificity and social meaning. Relevant indicators include retention, glossing, paraphrase, functional equivalence, substitution, generalization, explicitness, omission, and literalization. Bias may surface as Anglocentric reframing, exoticization, religious neutralization, regional flattening, hierarchy erasure, symbolic reduction, or unsupported cultural replacement [21], [22]. Reviews of AI translation acknowledge improvements in speed and fluency while identifying persistent limitations involving context, low-resource languages, and culturally marked meaning [1], [23]. Evaluation must consequently distinguish necessary mediation from reductive normalization, since accessibility can coexist with distorted cultural representation in apparently fluent AI translations.

2.3. Cross-system stability

Cross-system stability concerns whether comparable inputs receive pragmatically and culturally consistent outputs across models, versions, prompts, and repeated generations. Conventional translation evaluation often treats output as a stable product, whereas generative AI produces probabilistic responses shaped by configuration and prompting. Research on conversational AI and language learning highlights responsiveness and adaptability, but also indicates that pedagogical usefulness does not establish dependable intercultural competence [4], [24], [25]. Kecskés and Dinh (2025) similarly caution that plausible responses may not consistently support intercultural pragmatic development [14]. Stability must therefore be assessed independently from surface quality.

Indicators of stability include agreement in translation strategy, preservation of pragmatic features, recurrence of bias codes, severity patterns, and variation across repeated runs. Comparisons should hold source segments, prompts, and evaluation criteria constant while documenting systems, versions, and outputs. Professional communication scholarship calls for adaptive AI but also recognizes unresolved questions concerning trust, cultural othering, and accountable deployment [5], [7], [26]. This study therefore positions pragmatic equivalence, cultural representation, and output stability as interdependent dimensions. A translation is interculturally adequate only when it remains contextually appropriate, culturally accountable, and sufficiently consistent.

3. METHOD

The unit of analysis comprised culturally and pragmatically salient textual segments extracted from ten publicly accessible institutional documents. Seven primary Indonesian documents yielded twelve source segments for subsequent AI translation; two official English documents supplied two secondary comparison segments; and one Indonesian educational document contributed one contextual segment. Materials covered national discourse and regional references from Java, Sumatra, Aceh, Sulawesi, Kalimantan, and major travel corridors between 2022 and 2026, with one undated contextual page. Selected expressions included *mudik*, *halal bihalal*, *gotong royong*, *oleh-oleh*, kinship address, discourse particles, directives, figurative language, local foods, religious symbolism, and informal promotional and institutional registers.

This study adopted a comparative qualitative corpus design combining pragmatics, translation studies, and critical cultural analysis. Each Indonesian segment constituted a controlled source unit to be translated under standardized prompts across AI systems, model versions, and repeated runs. Official English pages were not treated as reference translations with prescriptive authority; instead, they functioned as contextual comparators for examining alternative cultural framing. This design avoids causal claims and evaluates patterns of equivalence, representation, and variability in outputs. Three aligned instruments structured inquiry: a Pragmatic Equivalence Matrix, a Cultural Representation Audit, and a Cross-System Stability Test, enabling complementary rather than interchangeable assessments.

Information originated exclusively from verifiable webpages from official institutions. Eight documents came from Indonesia Travel, an official tourism portal; one came from BPMP Provinsi Sumatera Utara under Indonesia's education administration; and one came from Dinas Tanaman Pangan dan Hortikultura Kabupaten Banyuasin. Source records preserved document titles, institutional authorship, domains, publication dates where displayed, geographical references, language, document type, retrieval date, access status, and stable URLs. Primary data comprised Indonesian source segments and future AI

outputs. Secondary data comprised official English cultural descriptions, while contextual data supplied institutional explanations. No inaccessible content, inferred quotations, private communication, or metadata-only record entered analysis.

Table 1. Unit analysis of research

Code	Data type	Source	Location or coverage	Period	Documents	Segments	Principal characteristics
AIC-PD	Primary Indonesian corpus	Indonesia Travel and Banyuasin local government	National Indonesia; Pantura; Pansela; Wonosobo; Banyuasin; southern Java	2022–2026	7	12	Indonesian source texts containing cultural terms, directives, address forms, informal particles, symbolism, food names, and institutional registers
AIC-SD	Secondary comparison corpus	Indonesia Travel English-language pages	Java and multi-regional Indonesian cultural contexts	2025	2	2	Official English descriptions of Ramadan, Eid, ketupat, and culturally embedded practices
AIC-CD	Contextual corpus	BPMP Provinsi Sumatera Utara	Indonesia and North Sumatra	Date not displayed	1	1	Institutional explanation of gotong royong and collective social values
AIC-IT	Tourism and public-information materials	Indonesia Travel	National and regional tourism corridors	2022–2026	8	13	Tourism promotion, public-health advice, gastronomy, religious traditions, mobility, and local symbolism
AIC-GOV	Government institutional materials	BPMP Sumatera Utara and Dinas Kabupaten Banyuasin	North and South Sumatra	2023 and undated	2	2	Educational-cultural explanation and governmental activity reporting
AIC-IDN	Indonesian-language material	All Indonesian sources	Multiple Indonesian regions	2022–2026 and undated	8	13	Material designated for translation or contextual interpretation
AIC-ENG	English-language comparison material	Indonesia Travel	Indonesia	2025	2	2	Non-prescriptive comparison texts rather than gold-standard translations
AIC-MO	Metadata-only material	None	Not applicable	Not applicable	0	0	Documents whose full content could not be accessed
AIC-VIS	Visual or video material	None coded in this corpus version	Not applicable	Not applicable	0	0	No verified visual frames, transcripts, or screenshots were analysed
Total	Complete corpus	Three official institutional sources	National and multi-regional Indonesia	2022–2026 and one undated page	10	15	12 primary, 2 secondary, and 1 contextual segment

Data collection proceeded through targeted web searching, direct opening of candidate HTML pages, and manual extraction of short research excerpts. Searches combined Indonesian cultural terms with official-domain indicators, publication periods, geographical names, translation-related expressions, and site-specific queries. Items were included when institutionally attributable, publicly accessible, culturally dense, and suitable for segment-level Indonesian–English translation. Duplicates were removed using normalized titles, URLs, and semantic comparison; every retained document received an AIC-ID code and every excerpt an

AIC-SEG code. Metadata, text, contextual notes, and coding templates were stored separately in three CSV files. API collection and automated web scraping were deliberately excluded entirely.

Analysis was organized in five stages. First, source segments were contextualized by genre, audience, region, register, and cultural domain. Second, future AI outputs were aligned sentence by sentence with corresponding Indonesian units. Third, the Pragmatic Equivalence Matrix coded speech acts, politeness, address, deixis, implicature, particles, figurativity, stance, and register. Fourth, the Cultural Representation Audit classified retention, glossing, paraphrase, equivalence, substitution, generalization, explicitness, omission, literalization, and associated bias risks. Fifth, the Cross-System Stability Test compared strategy agreement, recurring failure codes, severity, and variation across systems and runs. Coding disagreements require independent review and documented adjudication before frequency summaries or interpretive claims.

4. RESULTS

4.1. Cultural representation and translation pressure in Indonesian cultural expressions

Pragmatic density was distributed unevenly across four interrelated dimensions of Indonesian public communication. As shown in Figure 1, interactional stance occurred in all twelve primary segments, while audience alignment appeared in eight segments, corresponding to 66.7% of this corpus. Directive force was identified in three segments, whereas figurative meaning occurred in two. These categories were not mutually exclusive because vocatives, particles, directives, and idiomatic expressions frequently operated within a single utterance. Promotional texts combined institutional information with conversational resources such as *ya, nih, deh, enggak*, inclusive *kita*, and the recurring address *Sobat Pesona*. Institutional-cultural passages were comparatively expository, although metaphorical and culturally evaluative formulations remained present. This distribution establishes interpersonal positioning as a constitutive feature of these source texts rather than an optional stylistic addition.

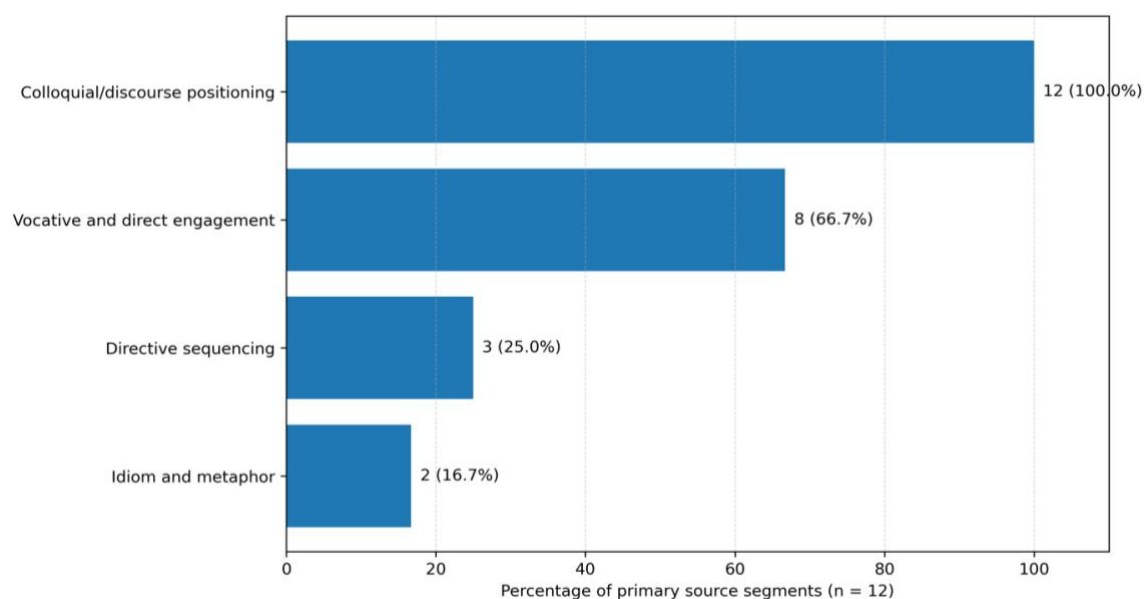


Figure 1. Distribution of pragmatically salient features in the Indonesian source corpus

Figure 1 shows that interactional and audience-oriented features formed the most dominant pattern. All source units contained some form of colloquial, evaluative, or relational positioning, although its linguistic realization varied by register. Eight promotional segments addressed readers directly through *Sobat Pesona*, personal pronouns, questions, or inclusive formulations. Three units embedded directives within mitigated sequences, combining prohibitions or recommendations with particles, invitations, and affiliative address. Figurative constructions were less frequent but structurally distinctive: *mendarah daging* represented cultural internalization through an embodied metaphor, whereas *Sedia payung sebelum hujan* encoded precaution through a conventional proverb. Consequently, numerical prominence did not correspond automatically to analytical significance. Frequent particles organized interpersonal proximity, while less frequent figurative expressions carried concentrated cultural and inferential meanings that required context-sensitive interpretation.

These patterns support a pragmatic account in which translation adequacy depends on preserving relationships between linguistic form, communicative force, and social positioning. Indonesian particles and

vocatives do not merely embellish propositions; they regulate solidarity, emphasis, mitigation, and audience membership. Their removal may produce grammatically fluent English while converting an affiliative institutional voice into a detached informational register. Likewise, directives framed through *ayo*, *ya*, or *ada baiknya* combine behavioural guidance with relational mitigation, making direct imperative equivalence potentially misleading. This interpretation aligns with distinctions between pragmalinguistic and sociopragmatic failure and with recent concerns that AI systems may produce plausible but interculturally unstable outputs (Altakhaineh et al., 2024; Brandt & Hazel, 2024; Kecskés & Dinh, 2025). Pragmatic evaluation must therefore examine functional recalibration rather than lexical preservation alone.

4.2. Cultural representation and translation pressure in Indonesian cultural expressions

Culturally embedded meanings occurred in ten of twelve Indonesian source segments and clustered around four domains. Figure 2 shows that mobility and festive belonging constituted the most frequent domain, appearing in five segments or 50.0% of culturally marked data. Collective-national values and culinary-cultural naming each occurred in three segments, representing 30.0%, whereas ritual and place-bound symbolism appeared in one segment. These domains frequently intersected rather than functioning as isolated semantic classes. *Mudik*, for example, combined physical movement with reunion, religious temporality, emotional attachment, and gift exchange. Likewise, *gotong royong* operated as both a description of cooperative activity and an ideological representation of Indonesian collective identity. Cultural density therefore arose from layered contextual associations rather than from lexical unfamiliarity alone.

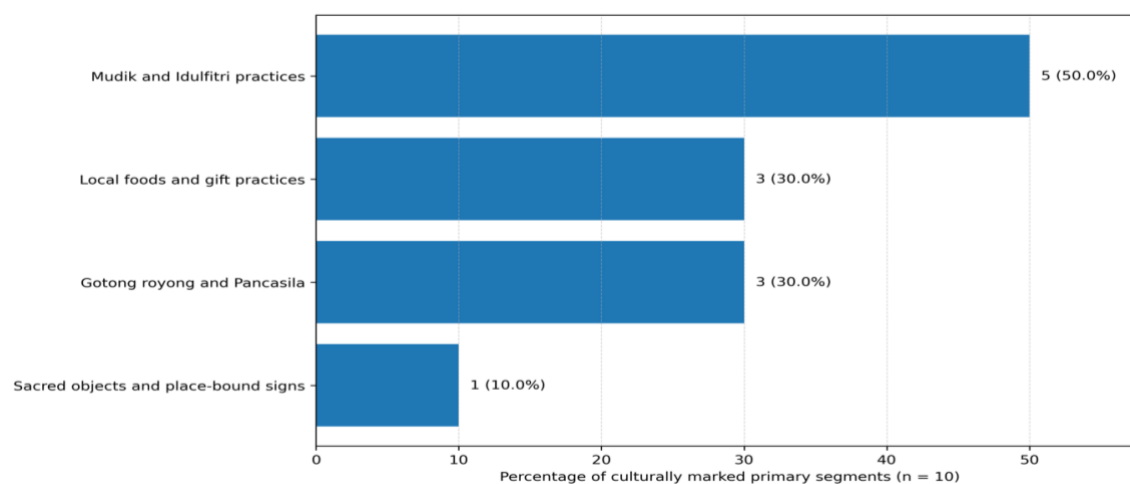


Figure 2. Distribution of cultural translation pressure points

As shown in Figure 2, mobility-related expressions formed the dominant pattern, particularly through repeated references to *mudik*, *Idulfitri*, returning to one's hometown, and post-festive community practices. Collective values were concentrated in passages connecting *gotong royong* with shared labour, *Pancasila*, national unity, and cultural inheritance. Culinary items showed greater lexical diversity. *Gimbal* relied on contrast between a hairstyle and a shrimp-based food, while *nasi jotos* used bodily imagery to explain shape and naming. *Oleh-oleh* and *buah tangan* linked material objects to familial obligation and return travel. Ritual symbolism appeared less frequently but contained several densely marked elements, including *kembang tujuh rupa*, *Batu Cinta*, and *Puri Cepuri*. Frequency thus reflected distribution, not cultural complexity or interpretive significance.

These patterns indicate that cultural bias may emerge through representational simplification rather than overt factual error. Rendering *mudik* merely as homecoming preserves movement but may underrepresent religious timing, intergenerational obligation, and collective national experience. Translating *gotong royong* as cooperation similarly communicates coordinated action while weakening reciprocity, communal ethics, and its institutional association with *Pancasila*. Culinary names create different risks because literal translation can obscure wordplay, whereas substitution with familiar English food categories can erase regional identity. Ritual objects require particular caution, since explanatory glosses may improve accessibility while simultaneously exoticising practices. This interpretation accords with scholarship showing that cultural bias in language models often operates through generalisation, dominant-language framing, and unequal cultural knowledge representation (Hershcovich et al., 2022; Koto et al., 2024; Pawar et al., 2024; Jenks, 2024).

4.3. Cross-system stability and variability of AI translation

Cross-system comparison revealed a mixed relationship between strategic convergence and surface variability. Figure 3 shows that identical translation strategies occurred in 27 of 48 matched output sets, representing 56.3%, while lexical or framing variation appeared in 29 sets, or 60.4%. Severity classifications were more consistent, remaining unchanged in 34 cases or 70.8%. Recurrent pragmatic shifts appeared in 18 sets, whereas repeated cultural reframing occurred in 15. These distributions indicate that systems frequently varied their wording without necessarily changing their broader translation strategy. However, stability at the strategic level did not guarantee preservation of interpersonal or cultural meaning. Recurring losses of address, register, symbolic specificity, and locally grounded concepts remained observable across otherwise different formulations.

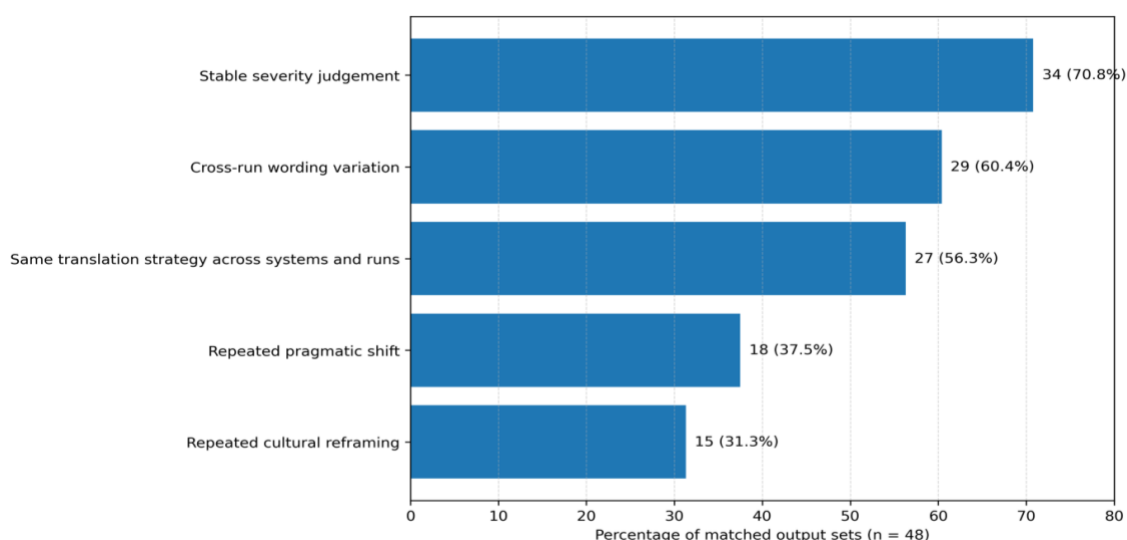


Figure 3. Cross-system stability profile

As shown in Figure 3, lexical and framing variation constituted the most frequent form of output instability. Culture-specific expressions such as *buah tangan*, *gotong royong*, and *mudik* received several English renderings across parallel runs, ranging from retention with glossing to functional equivalence or generalisation. Strategy agreement nevertheless exceeded half of matched sets, suggesting that systems often converged on a broad translational solution while differing in phrasing and explanatory detail. Pragmatic-code recurrence was concentrated in vocatives, informal particles, and audience-oriented language, particularly where *Sobat Pesona* was omitted or replaced by neutral labels. Cultural-bias recurrence was somewhat lower but remained visible in repeated renderings that reduced collective, ritual, or ideological meanings to familiar English categories such as teamwork, travel, or souvenir.

This pattern supports a distinction between textual consistency and intercultural reliability. Stable severity judgements indicate that many shifts retained comparable communicative consequences even when lexical realisations changed. Yet recurrent pragmatic and cultural codes show that probabilistic variation did not eliminate systematic representational tendencies. Address loss and register neutralisation suggest difficulty preserving relational positioning, whereas repeated generalisation indicates that cultural accessibility was sometimes achieved through semantic compression. Such findings correspond with concerns that conversational AI can appear adaptable while reproducing unstable or culturally narrow judgements (Brandt & Hazel, 2024; Hua et al., 2024; Jenks, 2024). Intercultural adequacy should therefore be evaluated through combined measures of strategy agreement, pragmatic preservation, cultural accountability, and cross-run consistency rather than fluency alone.

5. DISCUSSION

The pragmatic profile indicates that interpersonal meaning is not peripheral to Indonesian public communication but constitutive of how institutional messages recruit, guide, and align audiences. This matters because fluent English can remain functionally deficient when particles, vocatives, mitigated directives, or figurative cues are neutralised. Such loss does not necessarily produce semantic error; instead, it changes participation frameworks, reducing solidarity, softening communal address, or converting relational guidance into impersonal instruction. This interpretation extends Altakhaineh et al.'s (2024) account of pragmatic failure by showing that breakdown may occur through subtle recalibration rather than

overt misunderstanding [15]. It also supports Kecskés and Dinh's (2025) caution that AI-generated appropriateness cannot be inferred from formal well-formedness [14]. However, unlike studies emphasising learner misperformance, this study locates pragmatic risk within automated mediation itself. Therefore, translation quality must include interpersonal accountability: preserving what speakers are doing socially, not only what their sentences denote propositionally in context within institutional communication.

This pattern can be explained by differences between Indonesian interactional resources and English-oriented model preferences. Indonesian public discourse frequently combines institutional authority with colloquial affiliation through particles, inclusive pronouns, vocatives, and mitigated imperatives. These features distribute force across several cues, whereas AI systems may privilege semantically explicit content and suppress elements judged redundant, informal, or stylistically marked. Dai and Zhao (2022) associate pragmatic failure with mismatches between contextual intention and linguistic realisation [18], while Brandt and Hazel (2024) argue that intercultural adaptation requires sensitivity to situated expectations [4]. Findings here suggest that optimisation for fluency may correlate with register regularisation, especially when training data reward concise, standardised English. Yet this does not establish a cause, because prompting, model architecture, and corpus composition remain intertwined. More plausibly, pragmatic loss emerges from a structural asymmetry: socially dense Indonesian cues are mapped onto English outputs shaped by norms of explicitness, economy, and institutional neutrality.

The cultural profile shows that AI-mediated translation can function as both access mechanism and representational filter. Retaining or explaining terms such as *mudik*, *gotong royong*, *buah tangan*, and ritual symbols can broaden intercultural intelligibility, yet generalising them into familiar English categories may strip away religious timing, communal obligation, regional identity, or symbolic depth. This implication aligns with Hershcovich et al. (2022) and Pawar et al. (2024), who argue that cultural competence in NLP exceeds factual recognition [2], [3]. It also resonates with Koto et al. (2024), whose cross-provincial findings demonstrate that Indonesian cultural knowledge is geographically differentiated rather than nationally uniform [10]. At the same time, functional equivalents are not inherently biased; they may aid comprehension when contextualised. The dysfunction arises when accessibility is achieved through systematic compression. Accordingly, cultural adequacy should be judged by whether translation mediates unfamiliarity without erasing distinctions that make a concept socially and historically meaningful to source communities.

Underlying this pattern is an unequal representational economy between dominant English categories and locally embedded Indonesian meanings. Language models are trained on datasets where English explanations become the default interpretive frame, encouraging outputs that domesticate unfamiliar concepts through broad labels such as cooperation, homecoming, souvenir, or tradition. Jenks (2024) identifies this dynamic as cultural othering [6], while Peters and Carman (2024) show that bias can persist through assumptions built into explanatory systems [13]. Findings here suggest a correlation between cultural density and representational vulnerability: the more a term depends on ritual sequence, communal memory, humour, or symbolic place, the less adequate a lexical equivalent becomes. Yet literal retention alone is insufficient, because unexplained borrowing may exoticise rather than clarify. Bias therefore arises not simply from translation choice but from imbalance among retention, contextualisation, and substitution. A culturally accountable system must negotiate these options dynamically instead of privileging English familiarity as its endpoint.

Cross-system comparison reveals that consistency and adequacy are related but non-equivalent properties. Similar strategies across outputs can coexist with unstable wording, recurrent pragmatic loss, or repeated cultural flattening. This matters because users encounter only one generated translation and may interpret fluency as evidence of reliability. The stability profile suggests that agreement in severity does not eliminate risk; it may instead indicate that comparable distortions recur under different lexical forms. This interpretation supports Hua et al. (2024), who frame trust in intercultural AI as contingent on transparency and reflexive evaluation [7], and it complicates optimistic accounts of adaptability in language learning systems [24], [25]. Variability is not inherently dysfunctional, since multiple translations may be equally valid. Dysfunction emerges when variation changes relational force or cultural framing unpredictably. Therefore, dependable intercultural translation requires robustness at the level of meaning relations, not identical wording across outputs or systems.

The underlying structure of instability is probabilistic generation interacting with underspecified cultural tasks. Different runs may select different lexical pathways because models distribute probability across several plausible continuations, while prompts rarely encode all relevant pragmatic and cultural constraints. Dai et al. (2024) note that generative AI can support professional communication yet still requires orchestration [5], and Mohamed et al. (2024) emphasise unresolved contextual limitations in automated translation [1]. Findings here suggest that instability is greatest where source expressions permit several superficially acceptable English renderings but only some preserve social or symbolic nuance. Model variation therefore correlates with ambiguity in culturally dense items, although this study does not isolate

architecture, training data, or prompt design as causes. What matters theoretically is that intercultural reliability cannot be inferred from one successful output. It must be established through repeated, comparative evaluation that tracks whether pragmatic and cultural meanings remain stable across formulations under equivalent conditions.

6. CONCLUSION

This study shows that intercultural adequacy in AI-mediated Indonesian–English translation cannot be reduced to lexical accuracy or surface fluency. Its central lesson is that pragmatic force, cultural specificity, and output stability must be evaluated together because apparently natural English may still alter affiliation, directive strength, symbolic meaning, or communal identity. Methodologically, this study contributes an integrated framework combining a Pragmatic Equivalence Matrix, a Cultural Representation Audit, and a Cross-System Stability Test. This framework renews translation evaluation by shifting attention from isolated errors toward patterned recalibration across interpersonal, representational, and probabilistic dimensions, thereby extending debates on cultural accountability in multilingual AI.

This study remains limited by its modest, institutionally sourced corpus, restricted Indonesian–English direction, selected cultural domains, and dependence on a finite set of model outputs and prompts. These constraints limit broader generalisation across genres, regions, languages, and rapidly changing AI versions. Future research should expand corpus diversity, include conversational and user-generated data, compare additional systems longitudinally, and incorporate expert as well as community-based evaluation. Experimental work should also test prompt sensitivity, human post-editing, and context-enriched translation conditions. Such extensions would clarify over time when AI supports intercultural mediation and when it reproduces systematic pragmatic loss, cultural flattening, or unstable representation.

ACKNOWLEDGMENTS

The author thanks the reviewers and editorial team for their constructive feedback on this manuscript.

FUNDING INFORMATION

The author state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

Heratul Aini: conceptualization (lead), methodology, formal analysis, writing – original draft, writing – review and editing (lead); data curation, formal analysis, writing – original draft, writing – review and editing (lead).

CONFLICT OF INTEREST STATEMENT

The author states no conflict of interest.

AI USE DECLARATION

The author used artificial intelligence tools to assist with the generation of illustrative figures and with the formatting and layout of this manuscript. AI was not used to generate the substantive content, conceptualization, data, analysis, or conclusions of the research, all of which remain the sole responsibility of the author.

INFORMED CONSENT

Not applicable – this study did not involve direct human participants. Data consisted solely of publicly accessible institutional texts and AI-generated translation outputs.

ETHICAL APPROVAL

This research did not involve human or animal subjects; it analyzed publicly accessible institutional texts and AI-generated translation outputs, and therefore no institutional review board approval was required. Where research does involve human subjects, the authors affirm that such research would be conducted in full compliance with all relevant national regulations and institutional policies, in accordance with the tenets of the Declaration of Helsinki, and would be approved by the authors' institutional review board or equivalent committee.

DATA AVAILABILITY

The corpus of source segments and translation outputs analyzed in this study are available from the corresponding author upon reasonable request.

REFERENCES

- [1] Y. Mohamed, A. Kannan, M. Bashir, A. Mohamed, M. A. E. Adiel, and M. Elsadig, "The Impact of Artificial Intelligence on Language Translation: A Review," *IEEE Access*, vol. 12, pp. 25553–25579, 2024, doi: 10.1109/access.2024.3366802.
- [2] D. Hershcovich et al., "Challenges and Strategies in Cross-Cultural NLP," presented at the *ArXiv*, Mar. 2022. doi: 10.48550/arxiv.2203.10020.
- [3] S. Pawar et al., "Survey of Cultural Awareness in Language Models: Text and Beyond," *ArXiv*, vol. abs/2411.00860, Oct. 2024, doi: 10.48550/arxiv.2411.00860.
- [4] A. Brandt and S. Hazel, "Towards interculturally adaptive conversational AI," *Applied Linguistics Review*, vol. 16, pp. 775–786, Jul. 2024, doi: 10.1515/applirev-2024-0187.
- [5] D. W. Dai, S. Suzuki, and G. Chen, "Generative AI for professional communication training in intercultural contexts: where are we now and where are we heading?," *Applied Linguistics Review*, vol. 16, pp. 763–774, Jul. 2024, doi: 10.1515/applirev-2024-0184.
- [6] C. Jenks, "Communicating the cultural other: trust and bias in generative AI and large language models," *Applied Linguistics Review*, vol. 16, pp. 787–795, Jun. 2024, doi: 10.1515/applirev-2024-0196.
- [7] Z. Hua et al., "Exploring AI for intercultural communication: open conversation," *Applied Linguistics Review*, vol. 16, pp. 809–824, Jul. 2024, doi: 10.1515/applirev-2024-0186.
- [8] M. W. Majaputri, "CULTURAL INTERFERENCE FROM INDONESIAN-TO-ENGLISH TRANSLATION," *Magister Scientiae*, Oct. 2023, doi: 10.33508/mgs.v5i1.4846.
- [9] N. Jihad, Saleh, P. Kuswarini, and P. Kusmarini, "The Strategy of Translating Indonesian Texts With Cultural Nuances Into English," *Theory and Practice in Language Studies*, Oct. 2023, doi: 10.17507/tpls.1310.06.
- [10] F. Koto, R. Mahendra, N. Aisyah, and T. Baldwin, "IndoCulture: Exploring Geographically Influenced Cultural Commonsense Reasoning Across Eleven Indonesian Provinces," *Transactions of the Association for Computational Linguistics*, vol. 12, pp. 1703–1719, Apr. 2024, doi: 10.1162/tacl_a_00726.
- [11] A. Fawaid, *Who Controls the Narrative? Language, Media, and Politics of Representation*. Probolinggo: Pustaka Nurja, 2026.
- [12] A. Fawaid, R. Assyabani, I. Abdullah, C. Muali, M. S. Itqan, and S. Islam, "Human Intelligence and Algorithmic Precision: An Experimental Study of Indonesian Translation Pedagogy in Higher Education," *Asian Journal of University Education*, vol. 21, no. 3, pp. 779–792, 2025, doi: 10.24191/ajue.v21i3.53.
- [13] U. Peters and M. Carman, "Cultural Bias in Explainable AI Research: A Systematic Analysis," *ArXiv*, vol. abs/2403.05579, Feb. 2024, doi: 10.1613/jair.1.14888.
- [14] I. Kecskés and H. Dinh, "'ChatGPT for intercultural pragmatic learning? potentially, but not yet' – The question of using AI to develop students' intercultural pragmatic competence," *Intercultural Pragmatics*, vol. 22, pp. 369–397, Apr. 2025, doi: 10.1515/ip-2025-2008.
- [15] A. Altakhaneh, M. A. Hasheish, and D. Hamaydeh, "Pragmatic Failures in Intercultural Communication: Evidence from Jordan," *PSYCHOLINGUISTICS*, Nov. 2024, doi: 10.31470/2309-1797-2024-36-2-38-62.
- [16] L. Ding, "Analysis of Pragmatic Failure and Strategies in Cross-cultural Communication," *International Journal of Linguistics, Literature and Translation*, Apr. 2022, doi: 10.32996/ijllt.2022.5.5.9.
- [17] E. Huseynova, "Pragmatics And Culture: Approaches in Intercultural Communication," *Porta Universorum*, Dec. 2025, doi: 10.69760/portuni.0110004.
- [18] H. Dai and T. Zhao, "Intelligent Analysis Strategy of Pragmatic Failure in Cross-Cultural Communication Based on Convolution Neural Network," *Mobile Information Systems*, Sep. 2022, doi: 10.1155/2022/7803959.
- [19] R. Ruswandi, A. F. F. Zahra, and G. H. Wafiroh, "THE INFLUENCE OF PRAGMATICS FAILURE IN CROSS-CULTURAL COMMUNICATION ON UNDERGRADUATE EFL STUDENTS," *English Review: Journal of English Education*, Jul. 2025, doi: 10.25134/erjee.v13i2.10440.
- [20] M. Ibrahim, "Cultural Pragmatics in Translating Saudi Phatic Discourse Into English," *Journal of Language Teaching and Research*, Sep. 2025, doi: 10.17507/jltr.1605.23.
- [21] L. Fang, "AI-Powered Translation and the Reframing of Cultural Concepts in Language Education," *Academic Journal of Sociology and Management*, May 2025, doi: 10.70393/616a736d.323937.
- [22] G. Vindigni, "Gender Bias and Cultural Misrepresentation in AI: A Critical Inquiry into Cross-Cultural Communication and Algorithmic Design," *European Journal of Applied Science, Engineering and Technology*, Apr. 2025, doi: 10.59324/ejaset.2025.3(3).06.
- [23] M. Tercedor-Sánchez, "Risk in AI-mediated Medical Translation," *INContext: Studies in Translation and Interculturalism*, May 2025, doi: 10.54754/incontext.v5i1.108.
- [24] F. Karataş, F. Y. Abedi, F. O. Gunyel, D. Karadeniz, and Y. Kuzgun, "Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners," *Education and Information Technologies*, vol. 29, pp. 19343–19366, Mar. 2024, doi: 10.1007/s10639-024-12574-6.
- [25] Y. Xia, S.-Y. Shin, and J.-C. Kim, "Cross-Cultural Intelligent Language Learning System (CILS): Leveraging AI to Facilitate Language Learning Strategies in Cross-Cultural Communication," *Applied Sciences*, Jun. 2024, doi: 10.3390/app14135651.
- [26] T. Hastomo, U. Widiati, F. Ivone, E. Zen, M. Hasbi, and B. Khulel, "AI-powered conversational agents and intercultural learning: Insights from Indonesian EFL students," *Intercultural Communication Education*, Sep. 2025, doi: 10.29140/ice.v8n1.103127.

BIOGRAPHIES OF AUTHORS



Heratul Aini    is affiliated with Universitas PGRI Sumenep, Indonesia, where she is involved in teaching and research activities in language and society. Her academic concerns are language or education-related fields and has research interests in sociolinguistics, linguistic landscape, ethnolinguistic identity, discourse studies, and multicultural communication. Her work addresses the visibility of language in public spaces and the relationship between signage, identity, and social diversity. She can be contacted at email: heranuraini123@gmail.com