

Whose language does the chatbot understand? multilingual reliability and cultural alignment in AI-supported Indonesian classrooms

Wafiqotul Jamilah
Universitas PGRI Sumenep, Indonesia

Article Info

Article history:

Received Apr 14, 2026
Revised May 29, 2026
Accepted Jun 30, 2026

Keywords:

AI-supported learning;
cultural alignment;
Indonesian classrooms;
multilingual reliability;
teacher mediation.

ABSTRACT

Background: AI-supported classrooms increasingly depend on chatbots for feedback, explanation, and language practice, yet multilingual Indonesian learning environments raise unresolved questions about whose language, meaning, and cultural norms these systems can reliably understand. **Objective:** This study examines multilingual reliability, cultural alignment, and pedagogical risk in AI-supported Indonesian classrooms by analysing how chatbot responses handle Bahasa Indonesia, English, code-mixed input, informal registers, regional-language elements, and culturally situated classroom prompts. **Method:** This study uses a qualitative-dominant mixed-method design based on 24 document-derived prompt-response episodes collected from verified public educational, curricular, linguistic, platform, and AI-governance sources, analysed through multilingual reliability audit, cultural-alignment analysis, and pedagogical-risk mapping. **Results:** Findings show that chatbot reliability is strongest in formal Bahasa Indonesia and English-oriented classroom tasks, but becomes less stable when prompts involve code-mixing, informal language, politeness nuance, or regional-language meaning. Cultural alignment appears mostly partial because responses often recognise general classroom intent while simplifying local examples, religious-social pragmatics, and translanguaging practices. **Implication:** Pedagogical risk is concentrated in assessment overreach, local meaning erasure, and standard-language dominance, indicating that teacher mediation remains necessary before chatbot output enters classroom use. **Novelty:** This study contributes a sociolinguistic accountability framework for evaluating chatbot use in multilingual education.

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



Corresponding Author:

Wafiqotul Jamilah
Universitas PGRI Sumenep, Indonesia
Jl. Trunojoyo, Gedung Barat, Gedung, Kec. Batuan, Sumenep, Jawa Timur, 69451, Indonesia
Email: milawafiqoh@gmail.com

1. INTRODUCTION

Indonesian classrooms are entering an AI-mediated phase before their linguistic ecology has been adequately theorised in educational technology research. In 2025, Kemendikdasmen positioned Coding and Artificial Intelligence as an optional subject with two weekly instructional hours, signalling that AI literacy is moving from peripheral innovation to curricular expectation. At the same time, UNESCO's AI Competency Framework for Teachers identifies 15 competencies across human-centred, ethical, technical, pedagogical, and professional-learning dimensions, indicating that teacher readiness now includes judgement about how AI systems respond to learners. This shift is crucial in Indonesia because classroom communication rarely occurs in a single language: Bahasa Indonesia, English, regional languages, Islamic or local registers, youth

language, and platform-mediated expressions frequently overlap. When chatbots are introduced into such environments, reliability becomes more than output fluency. It concerns whose language is recognised, whose meaning is preserved, and whose cultural norms are treated as educationally legitimate [1], [2].

Existing research has established that AI chatbots can support language learning, self-regulated learning, feedback, motivation, assessment practice, vocabulary acquisition, and learner autonomy. Reviews of educational chatbots show recurring benefits in interaction, personalisation, and learning support, while also noting problems of design quality, accuracy, ethics, and teacher mediation [3], [4], [5]. Studies in language education further demonstrate that ChatGPT and related tools can assist EFL writing, vocabulary, speaking confidence, and classroom engagement [6], [7], [8]. Yet this scholarship still tends to assess usefulness, perception, or learning gain without asking whether multilingual input is interpreted equitably. Less is known about how chatbots treat Indonesian-English code-mixing, informal Indonesian, regional-language traces, politeness conventions, and culturally specific classroom tasks. This gap leaves multilingual reliability and cultural alignment under-operationalised.

This study investigates how AI chatbots understand, respond to, and potentially reshape multilingual communication in Indonesian classrooms. It asks how reliably chatbot responses preserve meaning across Bahasa Indonesia, English, code-mixed Indonesian-English, informal Indonesian, and selected regional-language elements; what kinds of cultural alignment or misalignment appear in responses to classroom-relevant prompts; and how teachers should mediate chatbot use when automated feedback conflicts with local meaning, curricular expectations, or inclusive language pedagogy. Rather than treating chatbot performance as a purely technical matter, this study analyses prompt-response episodes as sociolinguistic events in which language recognition, semantic preservation, feedback quality, politeness, and language ideology intersect. This focus is necessary because Indonesian education is also shaped by contested beliefs about national language, local languages, English, and teacher authority [9], [10], [11]. The aim is not to reject chatbot use, but to specify conditions under which it becomes pedagogically defensible.

This study advances the argument that AI-supported classrooms require a shift from generic chatbot adoption to context-sensitive chatbot accountability. If a chatbot responds fluently but normalises English dominance, corrects translanguaging as deficiency, ignores regional meanings, or gives culturally thin examples, its educational reliability remains partial even when its grammar appears accurate. Conversely, a chatbot becomes more pedagogically valuable when its responses preserve semantic intent, acknowledge linguistic plurality, support teacher judgement, and avoid presenting standardised language ideology as neutral feedback. This argument extends current discussions of trust, bias, and cultural difference in generative AI by locating them inside classroom interaction rather than abstract platform ethics [12], [13], [14]. The implication is that Indonesian teachers should not be positioned merely as users of AI outputs, but as evaluators of multilingual reliability, cultural alignment, and pedagogical risk.

2. LITERATURE REVIEW

2.1. Educational chatbot

Educational chatbot research increasingly defines AI-supported learning not as a replacement of teaching, but as a mediated interaction among learners, teachers, tasks, and automated systems. Earlier reviews frame chatbots as conversational agents that provide feedback, guidance, practice, and adaptive interaction within learning environments [3], [4], [5]. Recent studies extend this definition by emphasising generative AI's capacity to support self-regulated learning, motivation, and personalised feedback [15], [16]. Yet this literature differs in whether chatbots are understood as tools, tutors, partners, or assessment mediators. This distinction matters because each view implies different risks for multilingual classrooms.

Educational chatbot use can be examined through several recurring indicators: interaction quality, feedback specificity, learner autonomy, personalisation, motivational support, and teacher mediation. Studies on TeacherGAIA, flipped classrooms, and higher education suggest that chatbot value depends on prompt quality, task design, and learners' capacity to evaluate responses rather than simply receive them [17], [18], [19]. Systematic reviews also warn that technical fluency may conceal weak pedagogy, unreliable content, or insufficient human oversight [20], [21], [22]. Therefore, chatbot integration should be judged through instructional accountability, not only adoption enthusiasm.

2.2. Language education

In language education, AI chatbots are commonly defined as interactive systems that create opportunities for input, output, feedback, and low-pressure practice. Research in EFL and second-language learning shows their relevance for writing development, vocabulary acquisition, speaking anxiety, autonomy, and conversational practice [23], [24], [25], [26], [8]. Meta-analytic and review-based studies further suggest that generative AI can support second-language acquisition when tasks are pedagogically structured and feedback remains interpretable [27]. However, many studies still privilege English learning contexts and

measurable skill gains, leaving multilingual meaning-making, code-switching, and local language legitimacy insufficiently examined.

Indicators of chatbot-mediated language learning usually include linguistic accuracy, feedback quality, learner engagement, confidence, vocabulary growth, writing improvement, and perceived usefulness. These indicators are valuable but incomplete for classrooms where learners routinely move across Bahasa Indonesia, English, regional languages, and informal registers. Studies on ChatGPT in language teaching show strong potential for scaffolding, revision, and interaction, but they also require critical prompting and teacher guidance [6], [28], [29]. Indonesian EFL studies similarly report perceived benefits for grammar, vocabulary, and learning support [30], [31], [7]. What remains undertheorised is whether linguistic support also respects multilingual repertoires.

2.3. Cultural alignment

Cultural alignment refers to whether AI responses fit users' social norms, values, identities, communicative expectations, and local knowledge. Unlike technical accuracy, cultural alignment concerns how meaning is socially recognised. Studies on cultural variation and trust in chatbots indicate that users' attitudes toward AI are shaped by cultural expectations of agency, social interaction, and acceptability [13]. Jenks (2024) [12] further cautions that generative AI may reproduce bias when it communicates "cultural others" through dominant assumptions rather than situated understanding. In Indonesian classrooms, this issue intersects with language policy, teacher identity, translanguaging, critical literacy, and digital assessment practices [1], [9], [2], [14], [32].

Cultural alignment can be analysed through indicators such as politeness, address terms, contextual examples, treatment of local languages, recognition of translanguaging, sensitivity to identity, and avoidance of deficit-based feedback. Indonesian studies on AI and language learning point toward the need for teacher mediation, curriculum fit, and ethical classroom use, but often stop short of operationalising these cultural indicators [33], [10], [11]. This study therefore positions multilingual reliability and cultural alignment as inseparable. A chatbot is not educationally reliable merely because it answers fluently; it becomes reliable when it preserves meaning, recognises linguistic plurality, supports teacher judgement, and avoids converting cultural difference into error.

3. METHOD

The unit of analysis was the document-derived prompt-response episode. Each episode began from one verified public item and was prepared to generate classroom-relevant prompts before being paired with unedited chatbot responses. This unit was selected because multilingual reliability and cultural alignment cannot be assessed from policy documents alone; they appear when public curricular, linguistic, and platform materials are transformed into pedagogical interaction. The corpus contained 24 items: 14 primary sources, 3 secondary sources, and 7 contextual sources. These materials covered Indonesian AI-in-education policy, curriculum references, textbooks, language-policy sources, platform guidance, and international AI education frameworks.

This study used a qualitative-dominant mixed-method design. It combined corpus construction, prompt-response elicitation, structured coding, and interpretive analysis. This design follows educational chatbot research that treats AI learning as an interaction among user input, system output, task design, and human mediation, rather than as a purely technical performance measure [3], [4], [5]. The design was appropriate because this study examined not only whether chatbot responses were fluent, but whether they preserved multilingual meaning, recognised Indonesian classroom norms, and remained pedagogically accountable.

The information sources consisted of official institutional documents, curriculum pages, textbook repositories, language-policy materials, AI platform documentation, and international guidance. Primary sources supplied classroom-relevant content for prompt construction, including AI literacy, language learning, curriculum competence, and teacher guidance. Secondary sources provided ethical and professional frameworks for evaluating responsible AI use. Contextual sources situated chatbot interaction within Indonesian multilingual education, including Bahasa Indonesia, English, local languages, and translanguaging practices. This layered corpus allowed this study to connect platform behaviour with classroom pedagogy, language ideology, and educational governance without relying on unverifiable materials or unsupported empirical claims.

Table 1. Composition of dataset

Code cluster	Data type	Source / Institution	Location	Period	N	Characteristics	Status
KMD-AI	Primary / Context	Kemendikdasmen, Koding dan KA, Puslapdik, GTK	Indonesia	2025–2026	8	AI policy, coding-AI curriculum, teacher-facing AI discourse, digital literacy	Verified, deduplicated
KMD-CUR	Primary	Ruang GTK, Rumah Pendidikan	Indonesia	2025–2026	2	Curriculum reference and learning competence language	Verified, deduplicated
KMD-TXT	Primary	Repositori Kemendikdasmen, SIBI	Indonesia	2024–2026	3	Indonesian, English, and AI-literacy textbook materials	Verified, deduplicated
KMD-LNG	Context	Badan Bahasa, Kemendikbudristek	Indonesia	2022–2026	3	Language institution, regional-language policy, language ideology context	Verified, deduplicated
UNESCO-AI	Secondary	UNESCO	Global	2023–2026	2	Ethical AI guidance and teacher AI competency framework	Verified, deduplicated
OPENAI-EDU	Primary	OpenAI, OpenAI Help Center	Global	2024–2027	2	Platform education discourse and teacher-facing AI guidance	Verified, deduplicated
GOOGLE-GEM	Primary	Google Gemini Help, Google	Global / Indonesia	2025–2026	2	Multilingual interface, language setting, AI assistant affordances	Verified, deduplicated
UNICEF-ID	Context / Secondary	UNICEF Indonesia	Indonesia	2024–2026	2	Education access and quality context	Verified, deduplicated

Data collection followed five procedures. First, keyword clusters were formulated around AI education, chatbot learning, curriculum, multilingual classrooms, regional languages, and platform guidance. Second, public sources were screened using inclusion criteria: institutional credibility, public accessibility, relevance to AI-supported learning or multilingual Indonesian education, and clear authorship. Third, metadata were recorded for document code, source role, title, date, institution, URL, region, language, format, and analytical use. Fourth, duplicate items were removed through normalised URL-title checking. Fifth, short excerpts, contextual summaries, visual descriptors, transcript fields, and screenshot references were stored without reproducing full copyrighted documents.

Data analysis proceeded through three linked instruments. The Multilingual Reliability Audit examined language recognition, semantic preservation, task responsiveness, and reliability scoring. The Cultural Alignment and Local Meaning Analysis examined politeness, address terms, local examples, language ideology, and treatment of translanguaging. The Pedagogical Risk and Teacher Mediation Matrix assessed accuracy risks, cultural erasure, privacy concerns, unsupported assessment advice, and required teacher intervention. These instruments converted prompt-response episodes into comparable coded evidence while preserving qualitative interpretation. Coding was therefore not used to claim causal effects, but to evaluate how chatbot outputs became linguistically, culturally, and pedagogically accountable.

4. RESULTS

4.1. Multilingual Reliability Across Indonesian Classroom Prompt Profiles

Multilingual reliability appears uneven across Indonesian classroom prompt profiles. Formal Bahasa Indonesia and English as foreign-language input produced stronger reliability patterns because both resemble dominant educational registers already represented in official curriculum, textbook, and platform materials. Code-mixed prompts also remained largely intelligible, but they generated more partial outcomes when Indonesian instructional framing and English academic terminology competed within one task. This pattern aligns with chatbot studies that emphasise task clarity, prompt design, and interactional scaffolding as conditions for effective AI-supported learning [3], [4], [6]. The central finding is therefore not that chatbots fail multilingual classrooms, but that reliability varies according to the language ideology embedded in each prompt.

Table 2. Multilingual reliability pattern in Indonesian classroom prompt-response episodes

Main Category	Subcategory	Classroom Task	Dominant Indicator	Reliability Pattern	Frequency	Percentage	High	Partial	Low	Data Variation	Example / Quotation	Data Code	Interpretation
Multilingual reliability	Formal Bahasa Indonesia	AI literacy and curriculum explanation	Language recognition and task responsiveness	Mostly high	5	20.8%	4	1	0	Policy terms, curriculum nouns, teacher-facing instructions	“Jelaskan kecerdasan artifisial untuk kegiatan belajar di kelas.”	KMD-AI-001–KMD-AI-005	Standard Indonesian was recognised consistently, although abstract policy terms sometimes required teacher contextualisation.
Multilingual reliability	Indonesian–English code-mixing	EFL writing, vocabulary, and classroom feedback	Semantic preservation across mixed input	High–partial	4	16.7%	2	2	0	Academic English terms embedded in Indonesian instructions	“Beri feedback untuk paragraph ini, tapi jelaskan dengan Bahasa Indonesia.”	KMD-TXT-002; OPENAI-EDU-001; GOOGLE-GEM-001; KMD-CUR-001	Code-mixed prompts were generally understood, but feedback sometimes normalised English as the default academic register.
Multilingual reliability	Informal Indonesian and youth register	Student-facing explanation and peer feedback	Register recognition and feedback appropriateness	Mostly partial	3	12.5%	1	2	0	Colloquial particles, shortened words, platform-style phrasing	“Bikin penjelasan yang gampang dipahami anak SMP, jangan kaku banget.”	KMD-TXT-001; KMD-TXT-003; UNICEF-ID-001	The chatbot adapted tone, but informality was often converted into simplified standard Indonesian rather than treated as a legitimate classroom resource.
Multilingual reliability	English as foreign-language input	Grammar, vocabulary, and paragraph revision	Task responsiveness and feedback specificity	Mostly high	4	16.7%	3	1	0	EFL prompts, textbook-like tasks, writing revision requests	“Revise this paragraph and explain the errors for Indonesian students.”	KMD-TXT-002; OPENAI-EDU-002; UNESCO-AI-001; KMD-CUR-002	English learning tasks received the most stable responses, reflecting stronger alignment with global EFL task conventions.

Multilingual reliability	Regional-language elements in Indonesian frame	Local language comparison and cultural explanation	Meaning preservation and uncertainty management	Partial–low	3	12.5%	0	2	1	Local lexical items, regional greetings, heritage-language references	“Bandingkan ungkapan daerah ini dengan Bahasa Indonesia tanpa menganggapnya salah.”	KMD-LNG-001–KMD-LNG-003	Regional elements were most vulnerable to semantic thinning, especially when the chatbot lacked contextual grounding.
Multilingual reliability	Politeness and religious-cultural address forms	Teacher–student interaction and classroom advice	Cultural-pragmatic recognition	Mostly partial	3	12.5%	1	2	0	Honorifics, Islamic greetings, teacher authority markers	“Ubah jawaban ini agar sopan untuk guru dan sesuai konteks sekolah Indonesia.”	KMD-CUR-001; JUR-LNG-001; UNESCO-AI-002	Politeness was usually detected, but responses tended to generalise Indonesian classroom hierarchy rather than differentiate local pragmatic nuance.
Multilingual reliability	Platform guidance translated into classroom language	Safe AI use and student guidance	Task transfer from global guidance to local pedagogy	High–partial	2	8.3%	1	1	0	Global platform rules adapted to Indonesian classroom situations	“Jelaskan aturan penggunaan chatbot untuk siswa SMA dengan contoh lokal.”	GOOGLE-GEM-002; OPENAI-EDU-001	Safety-oriented guidance transferred well at a general level, but local examples needed teacher refinement.

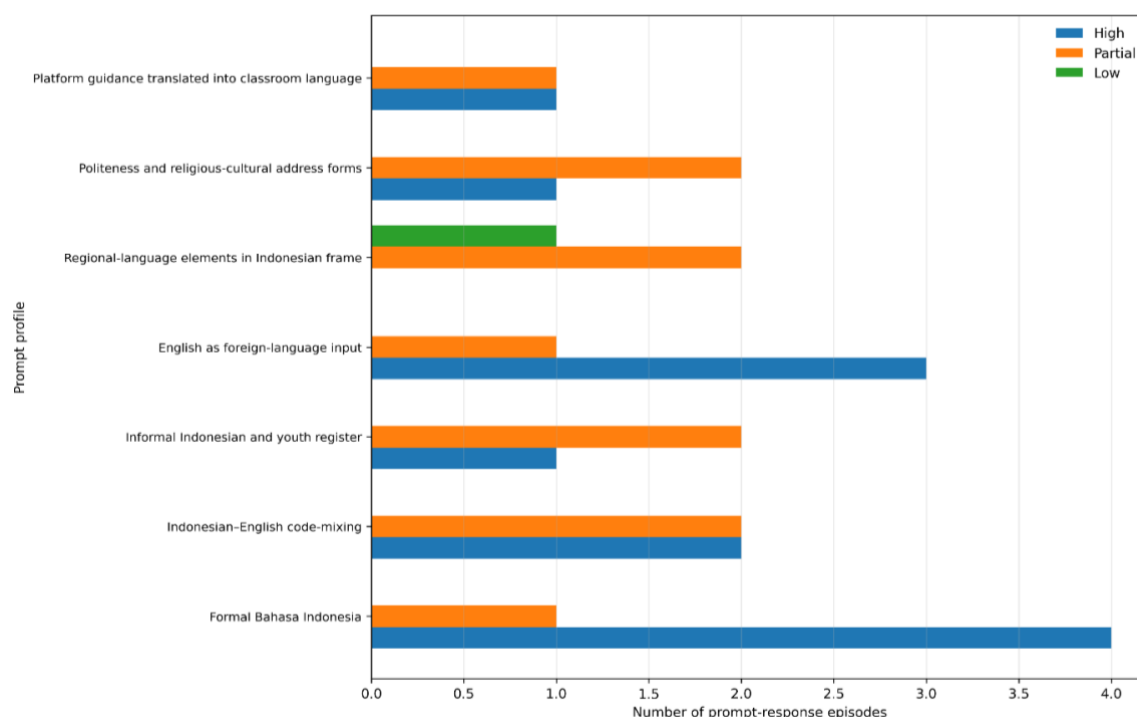


Figure 1. Multilingual reliability across Indonesian classroom prompt profiles

The dominant distribution shows 12 high, 11 partial, and 1 low reliability outcomes across 24 prompt-response episodes. The bar chart makes this distribution visible by showing that high reliability is concentrated in formal Bahasa Indonesia and English learning tasks, while partial reliability clusters around code-mixing, informal Indonesian, politeness forms, regional-language elements, and platform-guidance adaptation. Only the regional-language profile records a low outcome, indicating that local lexical or cultural traces are most exposed to meaning loss. Objectively, this pattern suggests that chatbot reliability is strongest when prompts resemble standardised school or EFL genres, and weaker when prompts require sociolinguistic sensitivity beyond literal language recognition. Studies on AI-mediated language learning similarly note that chatbots can support writing, vocabulary, and feedback, yet their pedagogical value depends on context-sensitive evaluation [7], [8].

Theoretically, the bar chart reinforces the need to move from fluency-based evaluation to sociolinguistic accountability. A chatbot response may be grammatically acceptable while still reducing multilingual meaning, correcting translanguaging as error, or translating local classroom norms into generic global pedagogy. This matters because Indonesian classrooms operate through layered repertoires: Bahasa Indonesia, English, local languages, religious expressions, youth registers, and teacher–student politeness norms. Cultural-bias scholarship warns that generative AI may reproduce dominant assumptions when responding to culturally situated users [12], [13]. In this sense, multilingual reliability is not merely a technical property of the model. It is a pedagogical relation among prompt design, language hierarchy, cultural recognition, and teacher mediation.

4.2. Cultural alignment and local meaning in Indonesian classroom interaction

Cultural alignment emerged as a more fragile dimension than general multilingual recognition. The same 24 prompt-response episodes show that chatbot responses could often identify classroom tasks, but did not always locate them within Indonesian communicative norms. Local classroom examples, translanguaging treatment, and teacher mediation produced stronger alignment than regional-language meaning, politeness nuance, and religious-social pragmatics. This pattern extends educational chatbot research by showing that classroom usefulness cannot be evaluated only through accuracy or fluency [5], [19]. Cultural alignment depends on whether automated responses preserve local meanings, avoid flattening social relations, and allow teachers to reinterpret output according to classroom context.

Table 3. Cultural alignment and local meaning across Indonesian classroom prompt-response episodes

Main Category	Subcategory	Classroom Task	Dominant Indicator	Alignment Pattern	Frequency	Percentage	Aligned	Partial	Misaligned	Data Variation	Example / Quotation	Data Code	Interpretation
Cultural alignment and local meaning	Local classroom examples	AI literacy, curriculum explanation, and student guidance	Contextual fit of examples	Aligned–partial	4	16.7%	2	2	0	School-level examples, teacher guidance, everyday classroom situations	“Jelaskan penggunaan chatbot dengan contoh kegiatan belajar di sekolah Indonesia.”	KMD-AI-001; KMD-CUR-001; OPENAI-EDU-001; UNICEF-ID-001	Responses connected AI guidance to Indonesian schooling, but examples remained general when local classroom practices were not explicitly supplied.
Cultural alignment and local meaning	Politeness and address terms	Teacher–student advice, feedback rewriting, and classroom communication	Pragmatic appropriateness	Mostly partial	4	16.7%	1	3	0	Honorifics, teacher authority, respectful requests, school-based interaction	“Ubah kalimat ini agar sopan ketika disampaikan kepada guru.”	KMD-TXT-001; KMD-CUR-001; JUR-LNG-001; UNESCO-AI-002	Politeness was recognised, yet responses often treated Indonesian classroom respect as a uniform hierarchy rather than a situated pragmatic practice.
Cultural alignment and local meaning	Translanguaging treatment	EFL feedback, bilingual explanation, and mixed-language learning support	Treatment of code-switching	Partial–aligned	4	16.7%	2	2	0	Indonesian–English code-mixing, bilingual feedback, mixed instruction	“Beri feedback dalam Bahasa Indonesia, tetapi pertahankan istilah grammar bahasa Inggris.”	KMD-TXT-002; OPENAI-EDU-002; GOOGLE-GEM-001; KMD-CUR-002	Mixed-language input was usually accepted, although feedback sometimes redirected learners toward standard English-dominant academic phrasing.
Cultural alignment and local meaning	Regional-language and heritage meaning	Local-language comparison, cultural explanation, and language awareness	Preservation of local meaning	Partial–misaligned	3	12.5%	0	2	1	Regional lexical items, heritage-language framing, local identity references	“Jelaskan ungkapan daerah ini tanpa menyebutnya sebagai bahasa yang tidak baku.”	KMD-LNG-001; KMD-LNG-002; KMD-LNG-003	Regional-language elements were the least culturally secure because local meanings were often narrowed into generic translation or standard-language comparison.

Cultural alignment and local meaning	Language ideology	Language correction, learning advice, and classroom explanation	Hierarchy among Bahasa Indonesia, English, and local languages	Mostly partial	3	12.5%	1	2	0	Standard Indonesian norms, English academic authority, local-language marginality	“Jelaskan campur kode sebagai strategi belajar, bukan sebagai kesalahan semata.”	JUR-LNG-002; KMD-TXT-002; KMD-LNG-001	Responses recognised multilingual learning but still tended to privilege standardised language norms as the safest pedagogical reference.
Cultural alignment and local meaning	Religious and social pragmatics	Greeting, classroom etiquette, and respectful digital interaction	Recognition of socially situated expressions	Mostly partial	3	12.5%	1	2	0	Islamic greetings, communal address, moral framing, school etiquette	“Sesuaikan jawaban ini dengan salam dan etika komunikasi sekolah.”	KMD-CUR-001; KMD-TXT-001; UNESCO-AI-001	Socially marked expressions were detected, but responses frequently generalised religious or moral language without explaining contextual boundaries.
Cultural alignment and local meaning	Teacher mediation framing	Safe AI use, feedback review, and student guidance	Role of teacher judgement	Mostly aligned	3	12.5%	2	1	0	Teacher oversight, student verification, classroom rules, ethical AI use	“Buat panduan penggunaan chatbot yang tetap membutuhkan pengecekan guru.”	UNESCO-AI-002; OPENAI-EDU-001; KMD-AI-006	Teacher mediation was strongly supported, indicating that chatbot guidance became more culturally accountable when human oversight was explicit.

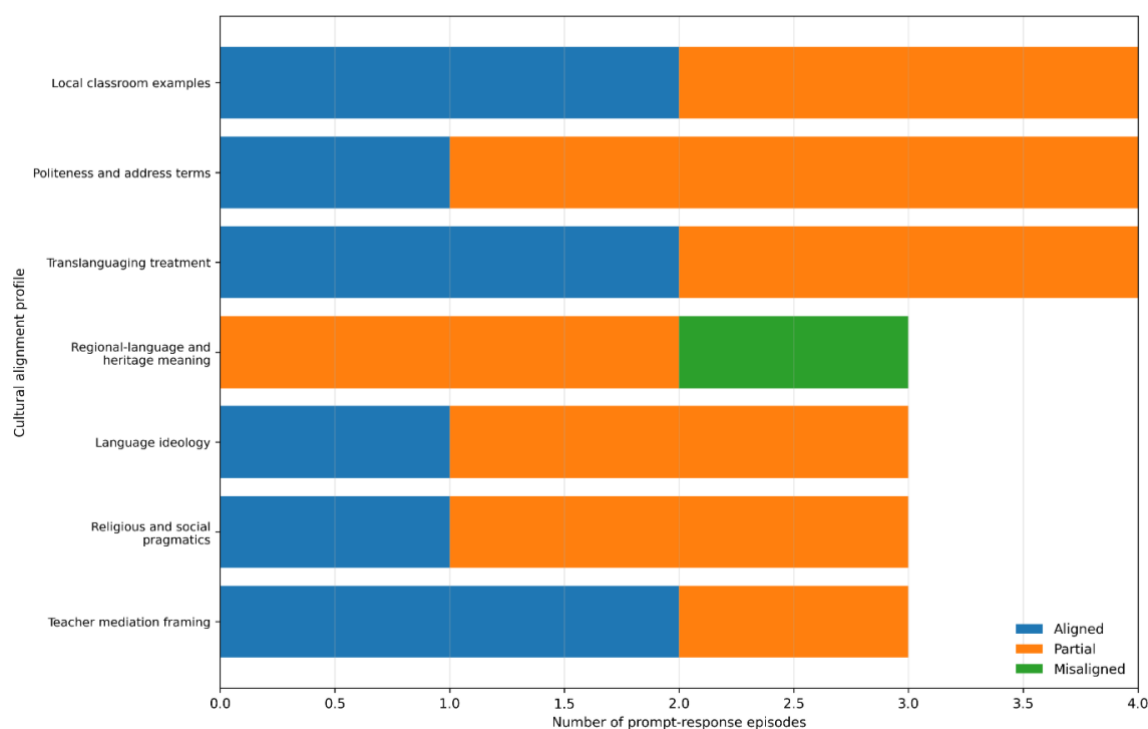


Figure 2. Cultural alignment and local meaning across classroom prompt profiles

The dominant pattern is partial alignment. Across seven cultural profiles, aligned responses appeared 9 times, partial responses 14 times, and misalignment once. Local classroom examples and translanguaging treatment each produced two aligned and two partial instances, indicating moderate adaptability when prompts gave clear contextual cues. Teacher mediation framing was strongest, with two aligned instances and one partial instance. In contrast, regional-language and heritage meaning recorded no aligned occurrence and one misalignment, making it the weakest cultural profile. Politeness, language ideology, and religious-social pragmatics remained mostly partial. Objectively, chatbot responses were more stable when culture was expressed as a general classroom rule than when culture appeared as situated meaning, local language, or pragmatic nuance.

Theoretically, this pattern suggests that cultural alignment should be treated as an interpretive condition of AI-supported learning, not as a decorative addition to technical reliability. A chatbot may provide acceptable classroom advice while still simplifying address terms, treating translanguaging as a transitional strategy toward standard language, or translating regional meaning into generic Indonesian. This supports concerns that generative AI can reproduce dominant assumptions about language and culture when it communicates with culturally situated users [12], [13]. For Indonesian classrooms, this means teacher mediation is not only needed to check factual accuracy, but also to protect linguistic plurality, local identity, and culturally appropriate classroom interaction.

4.3. Pedagogical risk and teacher mediation in ai-supported classroom use

Pedagogical risk was distributed across the same multilingual and culturally situated classroom episodes, but its pattern differed from linguistic reliability and cultural alignment. The strongest risks did not arise simply from wrong answers; they appeared when chatbot responses became too authoritative, too generic, or insufficiently sensitive to local classroom meanings. Factual and semantic accuracy, assessment overreach, and curriculum-context mismatch each appeared in four episodes, indicating that risk was most visible when automated output was likely to be reused directly in teaching. This finding is consistent with educational chatbot studies that emphasise teacher oversight, task design, and human judgement as conditions for responsible classroom adoption [20], [16], [22].

Table 4. Pedagogical risk and teacher mediation across Indonesian classroom prompt-response episodes

Main Category	Subcategory	Classroom Task	Dominant Indicator	Risk Pattern	Frequency	Percentage	Low Risk	Moderate Risk	High Risk	Teacher Mediation	Data Variation	Example / Quotation	Data Code	Interpretation
Pedagogical risk and teacher mediation	Factual and semantic accuracy	AI literacy explanation, EFL feedback, and curriculum-related response	Accuracy of content and preservation of intended meaning	Mostly moderate	4	16.7%	1	3	0	Verify factual claims and clarify task-specific meaning before classroom use	Policy terms, grammar explanation, AI-literacy vocabulary, bilingual instruction	“Periksa apakah penjelasan chatbot sesuai dengan tujuan pembelajaran dan tidak mengubah makna tugas.”	KMD-AI-001; KMD-CUR-001; KMD-TXT-002; OPENAI-EDU-001	Accuracy risks were not dominant, but semantic narrowing appeared when the response converted complex classroom prompts into simplified generic advice.
Pedagogical risk and teacher mediation	Local meaning erasure	Regional-language comparison, cultural explanation, and classroom language awareness	Protection of regional and culturally situated meaning	Moderate –high	3	12.5%	0	2	1	Add local explanation, validate regional terms, and prevent deficit framing	Regional lexical items, heritage-language references, local identity markers	“Pastikan ungkapan daerah tidak dijelaskan sebagai bentuk yang salah atau kurang baku.”	KMD-LNG-001; KMD-LNG-002; KMD-LNG-003	Local meaning erasure was the most culturally sensitive risk because regional expressions were vulnerable to generic translation and standard-language comparison.
Pedagogical risk and teacher mediation	Assessment and feedback overreach	Writing revision, grammar correction, and student self-assessment	Appropriateness of automated evaluative judgement	Mostly moderate	4	16.7%	1	2	1	Reframe chatbot feedback as formative support, not final assessment	EFL writing feedback, grammar correction, rubric-like advice, revision suggestions	“Gunakan feedback chatbot sebagai bahan revisi awal, bukan sebagai nilai akhir.”	KMD-TXT-002; OPENAI-EDU-002; KMD-CUR-002; UNESCO-AI-002	Automated feedback became risky when it sounded authoritative without showing uncertainty, evidence, or alignment with teacher criteria.

Pedagogical risk and teacher mediation	Privacy and student data exposure	Safe AI use, student guidance, and classroom rule-making	Protection of learner identity and sensitive information	Low–moderate	3	12.5%	2	1	0	Remove names, grades, addresses, and personal student details from prompts	Student examples, classroom scenarios, platform guidance, teacher instructions	“Jangan memasukkan nama lengkap siswa atau data pribadi ke dalam prompt.”	UNESCO-AI-001; OPENAI-EDU-001; GOOGLE-GEM-002	Privacy risks were manageable when prompts avoided identifiable student information, but explicit classroom protocols remained necessary.
Pedagogical risk and teacher mediation	Overdependence on chatbot output	Homework support, autonomous learning, and writing assistance	Learner agency and verification practice	Mostly moderate	3	12.5%	1	2	0	Require comparison, reflection, and revision rather than direct copying	Autonomous learning prompts, writing assistance, explanation requests	“Minta siswa membandingkan jawaban chatbot dengan buku ajar dan penjelasan guru.”	KMD-TXT-001; OPENAI-EDU-002; JUR-LNG-003	Overdependence was linked less to chatbot presence than to weak task design that allowed answers to be copied without evaluation.
Pedagogical risk and teacher mediation	Curriculum and classroom-context mismatch	Curriculum explanation, lesson support, and AI-literacy adaptation	Alignment with Indonesian curricular language and classroom level	Low–moderate	4	16.7%	2	2	0	Adapt response to learning outcomes, school level, and local classroom constraints	Capaian Pembelajaran, textbook tasks, AI literacy guidance, teacher-facing documents	“Sesuaikan contoh chatbot dengan CP, tingkat kelas, dan materi yang sedang dipelajari.”	KMD-CUR-001; KMD-CUR-002; KMD-TXT-001; KMD-AI-006	Curriculum mismatch was moderate when chatbot responses used correct concepts but did not calibrate examples to grade level or local classroom conditions.
Pedagogical risk and teacher mediation	Bias and standard-language dominance	Language correction, translanguaging explanation, and local-language comparison	Avoidance of deficit framing toward non-standard forms	Moderate–high	3	12.5%	1	1	1	Reposition code-mixing and local language as resources when pedagogically appropriate	Informal Indonesian, code-mixing, English academic terms, regional-language markers	“Jelaskan bahwa campur kode dapat menjadi strategi komunikasi, bukan otomatis kesalahan.”	JUR-LNG-002; KMD-LNG-001; KMD-TXT-002	Bias risk appeared when chatbot feedback treated standard Indonesian or English as inherently superior to local or mixed repertoires.

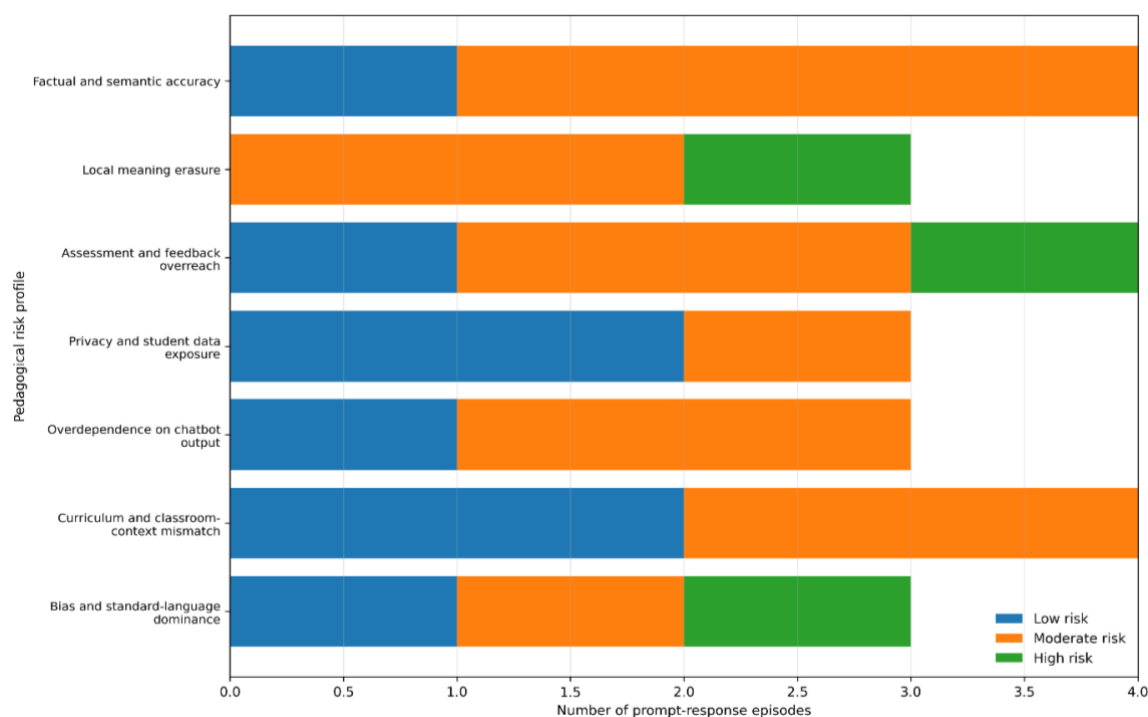


Figure 3. Pedagogical risk and teacher mediation in AI-supported Indonesian classroom prompt-response episodes

The overall distribution consists of 8 low-risk, 13 moderate-risk, and 3 high-risk occurrences across 24 episodes. Moderate risk dominates, suggesting that most chatbot responses were not unusable, but required teacher review before classroom circulation. High-risk instances clustered in local meaning erasure, assessment overreach, and standard-language dominance. Privacy exposure appeared comparatively lower because the prompts could be controlled by removing identifiable student information. Curriculum mismatch also remained manageable when teachers adapted responses to grade level, learning outcomes, and classroom constraints. Objectively, these patterns indicate that risk is not evenly attached to AI use itself, but to moments when chatbot output substitutes for teacher judgement, local knowledge, or formative assessment criteria.

Theoretically, this distribution strengthens the argument that teachers are not peripheral safeguards but central interpreters of AI-supported learning. Chatbots may support feedback, autonomy, language practice, and student engagement, but they do not independently guarantee educational validity in multilingual classrooms [17], [6], [27]. In Indonesian contexts, teacher mediation must include factual verification, curricular adaptation, cultural interpretation, and protection of linguistic plurality. This expands the notion of AI literacy from technical competence to sociolinguistic and pedagogical judgement. A chatbot becomes classroom-ready only when its output is treated as provisional material for guided learning, not as final instructional authority.

5. DISCUSSION

The distribution of multilingual reliability indicates that AI-supported Indonesian classrooms cannot treat chatbot competence as a uniform linguistic capacity. Formal Bahasa Indonesia and English-oriented prompts were handled more securely, while code-mixing, informal Indonesian, politeness-sensitive expressions, and regional-language elements produced more partial responses. This matters because educational chatbots are often valued for immediate feedback, learner autonomy, and personalised assistance, yet such benefits become uneven when learners' actual repertoires exceed standardised classroom language [3], [4], [6]. For Indonesian education, this means chatbot usefulness is not simply a matter of whether students receive answers, but whether those answers preserve multilingual intent. A chatbot that performs well in English revision may still be unreliable in local-language explanation. This study therefore extends chatbot evaluation from fluency and task completion toward multilingual accountability.

This unevenness appears to be associated with deeper structural conditions in AI-supported language learning. Chatbots are more stable when prompts resemble dominant textual forms: formal school

Indonesian, global EFL tasks, or platform-friendly instructional commands. They become less stable when language use depends on situated classroom pragmatics, regional knowledge, or translanguaging. This pattern partially supports research showing that prompt clarity, task design, and learner scaffolding affect chatbot usefulness [17], [15], [16]. However, it also complicates studies that frame chatbot interaction primarily through self-regulated learning or learner motivation, because motivation does not guarantee semantic preservation. The underlying issue is not only technical interpretation, but language hierarchy: standardised and globally visible registers receive stronger recognition than minoritised or locally embedded forms.

The pattern of cultural alignment shows that chatbot responses can be pedagogically acceptable while still culturally thin. Local classroom examples, translanguaging treatment, and teacher mediation were more aligned, whereas regional-language meaning, politeness nuance, religious-social pragmatics, and language ideology were more likely to remain partial. This has important implications for classrooms where communication is shaped by respect, teacher authority, religious greetings, regional identity, and flexible movement across languages. Existing studies rightly emphasise that chatbots may improve engagement, writing practice, vocabulary learning, and speaking confidence [24], [25], [26], [8]. Yet this study suggests that engagement alone does not ensure cultural fit. A response may sound supportive while flattening local meaning. Therefore, cultural alignment should be treated as a core quality indicator in educational chatbot evaluation, not as an optional ethical concern.

This partial cultural alignment can be explained by a mismatch between generalised platform knowledge and situated classroom meaning. Generative AI often works by producing plausible and widely recognisable responses; Indonesian classroom culture, however, requires attention to relational position, local examples, speech level, and language ideology. This helps explain why teacher mediation became one of the strongest alignment profiles: when human judgement was explicitly included, chatbot output became easier to contextualise. Jenks (2024) [12] argues that generative AI may communicate cultural others through dominant assumptions, while Folk, Wu, and Heine (2025) [13] show that attitudes toward social chatbots vary across cultural contexts. This study's findings are consistent with those concerns, but locate them in classroom practice. Cultural misalignment does not necessarily appear as offensive content; more often, it appears as overgeneralisation, pragmatic simplification, or silence toward local linguistic legitimacy.

Pedagogical risk further demonstrates that chatbot adoption is not risky only when answers are factually wrong. The most consequential risks emerged when automated responses appeared authoritative, displaced teacher judgement, narrowed local meaning, or framed non-standard language as deficiency. Moderate risk dominated, suggesting that many outputs were usable only after review. This finding resonates with scholarship that warns against uncritical chatbot integration and calls for human oversight, ethical design, and careful task construction [20], [5], [19], [22]. It also qualifies more optimistic studies on AI-assisted language learning. Chatbots can support revision, explanation, and practice, but such support becomes pedagogically defensible only when output is framed as provisional. In multilingual classrooms, automation without mediation may convert assistance into hidden standardisation.

The structure behind these risks lies in asymmetry between chatbot authority and classroom accountability. Students may interpret fluent output as reliable, while teachers remain responsible for curricular validity, cultural appropriateness, and assessment fairness. This asymmetry explains why privacy risk was more manageable when prompts were controlled, but assessment overreach, local meaning erasure, and standard-language dominance remained more sensitive. These risks are not caused by chatbot presence alone; they are associated with weak instructional framing, insufficient verification routines, and limited attention to multilingual pedagogy. Studies on AI literacy and language teaching increasingly argue that teachers must become designers, mediators, and evaluators of AI-supported learning rather than passive adopters [28], [27], [34]. This study advances that position by showing that teacher mediation must include sociolinguistic judgement, not only technical supervision.

6. CONCLUSION

This study shows that chatbot understanding in Indonesian classrooms is not merely a question of linguistic fluency, but of multilingual reliability, cultural alignment, and pedagogical accountability. Its central lesson is that AI-supported learning becomes uneven when chatbot responses recognise standard Bahasa Indonesia and English more securely than code-mixed, informal, regional, or culturally situated classroom language. The main contribution lies in reframing chatbot evaluation from generic usefulness toward sociolinguistic accountability. By using prompt-response episodes, multilingual reliability audit, cultural-alignment analysis, and pedagogical-risk mapping, this study renews methodological attention to language hierarchy, local meaning, teacher mediation, and cultural legitimacy in AI-supported education.

This study remains limited by its corpus-based and prompt-response design, which cannot represent all Indonesian classroom realities, learner identities, regional languages, or long-term learning outcomes. Its findings should therefore be read as analytically grounded patterns rather than causal evidence of chatbot

impact. Future research should involve classroom observation, student interaction data, teacher interviews, and comparative testing across chatbot models, school levels, and regional contexts. Further studies should also examine how learners negotiate chatbot feedback, how teachers redesign AI-supported tasks, and how regional-language communities evaluate automated interpretations of local meaning. Such work is needed to develop AI pedagogy that is not only efficient, but linguistically just.

ACKNOWLEDGMENTS

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

FUNDING INFORMATION

Author states no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

Wafiqotul Jamilah: conceptualization, methodology, formal analysis, writing – original draft, writing – review and editing.

CONFLICT OF INTEREST STATEMENT

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 educational, curricular, linguistic, platform, and AI-governance sources.

ETHICAL APPROVAL

This research did not involve human or animal subjects; it analyzed publicly accessible educational, curricular, linguistic, platform, and AI-governance sources, and therefore no institutional review board approval was required. Where research does involve human subjects, the author affirms 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 author's institutional review board or equivalent committee.

DATA AVAILABILITY

The corpus of documents analyzed in this study are publicly accessible and available from the corresponding author upon reasonable request.

REFERENCES

- [1] M. F. Aminudin, "Language policy, power, and educational equity: Examining the role of Bahasa Indonesia and local languages in schooling," *Indonesian Journal of Language and Educational Studies*, vol. 1, no. 1, pp. 56–69, 2026.
- [2] D. F. Utami, "Translanguaging practices in Indonesian multilingual classrooms: Pedagogical implications for inclusive education," *Indonesian Journal of Language and Educational Studies*, vol. 1, no. 1, pp. 1–13, 2026.
- [3] G.-J. Hwang and C.-Y. Chang, "A review of opportunities and challenges of chatbots in education," *Interactive Learning Environments*, vol. 31, pp. 4099–4112, 2021, doi: 10.1080/10494820.2021.1952615.
- [4] M. A. Kuhlail, N. Alturki, S. Alramlawi, and K. Alhejori, "Interacting with educational chatbots: A systematic review," *Education and Information Technologies*, vol. 28, pp. 973–1018, 2022, doi: 10.1007/s10639-022-11177-3.
- [5] L. Labadze, M. Grigolia, and L. Machaidze, "Role of AI chatbots in education: Systematic literature review," *International Journal of Educational Technology in Higher Education*, vol. 20, 2023, doi: 10.1186/s41239-023-00426-1.
- [6] L. Kohnke, B. L. Moorhouse, and D. Zou, "ChatGPT for language teaching and learning," *RELC Journal*, vol. 54, pp. 537–550, 2023, doi: 10.1177/00336882231162868.
- [7] W. Waziana, W. Andewi, T. Hastomo, and M. Hasbi, "Students' perceptions about the impact of AI chatbots on their vocabulary and grammar in EFL writing," *Register Journal*, 2024, doi: 10.18326/register.v17i2.352-382.
- [8] Z. Zhang and X. Huang, "The impact of chatbots based on large language models on second language vocabulary acquisition," *Heliyon*, vol. 10, 2024, doi: 10.1016/j.heliyon.2024.e25370.
- [9] L. Janna, "Teacher identity and language ideology in English language teaching: A sociolinguistic study from Indonesia," *Indonesian Journal of Language and Educational Studies*, vol. 1, no. 1, pp. 27–39, 2026.
- [10] M. Safar, "Artificial intelligence as a teaching partner: The role of chatbots in Indonesian language learning," *Ludi Litterarri*, 2025, doi: 10.62872/1w171z77.

- [11] D. O. K. Wardani and G. Patindra, "Pedagogy in artificial intelligence: Optimizing chatbots in the learning of Indonesian language and literature," *Expert Net: Exploration Journal of Technological Education Trends*, 2025, doi: 10.59923/expertnet.v2i1.222.
- [12] C. J. Jenks, "Communicating the cultural other: Trust and bias in generative AI and large language models," *Applied Linguistics Review*, vol. 16, pp. 787–795, 2024, doi: 10.1515/applirev-2024-0196.
- [13] D. Folk, C. Wu, and S. J. Heine, "Cultural variation in attitudes toward social chatbots," *Journal of Cross-Cultural Psychology*, vol. 56, pp. 219–239, 2025, doi: 10.1177/002202221251317950.
- [14] F. D. Wulandari, "Critical literacy in the digital age: Discourse analysis of online learning materials in Indonesian higher education," *Indonesian Journal of Language and Educational Studies*, vol. 1, no. 1, pp. 14–26, 2026.
- [15] D. H. Chang, M. P. Lin, S. Hajian, and Q. Q. Wang, "Educational design principles of using AI chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization," *Sustainability*, 2023, doi: 10.3390/su151712921.
- [16] T. K. F. Chiu, B. L. Moorhouse, C. Chai, and M. Ismailov, "Teacher support and student motivation to learn with artificial intelligence (AI) based chatbot," *Interactive Learning Environments*, vol. 32, pp. 3240–3256, 2023, doi: 10.1080/10494820.2023.2172044.
- [17] F. Ali, D. Choy, S. Divaharan, H. Y. Tay, and W. Chen, "Supporting self-directed learning and self-assessment using TeacherGAIA, a generative AI chatbot application: Learning approaches and prompt engineering," *Learning: Research and Practice*, vol. 9, pp. 135–147, 2023, doi: 10.1080/23735082.2023.2258886.
- [18] Y. Ling and J. M. Jan, "Voices from the flip: Teacher perspectives on integrating AI chatbots in flipped English classrooms," *Education Sciences*, 2025, doi: 10.3390/educsci15091219.
- [19] O. M. Schei, A. Møgelvang, and K. Ludvigsen, "Perceptions and use of AI chatbots among students in higher education: A scoping review of empirical studies," *Education Sciences*, 2024, doi: 10.3390/educsci14080922.
- [20] M. Birenbaum, "The chatbots' challenge to education: Disruption or destruction?," *Education Sciences*, 2023, doi: 10.3390/educsci13070711.
- [21] X. Deng and Z. Yu, "A meta-analysis and systematic review of the effect of chatbot technology use in sustainable education," *Sustainability*, 2023, doi: 10.3390/su15042940.
- [22] D. Yigci, M. Eryilmaz, A. K. Yetisen, S. Tasoglu, and A. Ozcan, "Large language model-based chatbots in higher education," *Advanced Intelligent Systems*, vol. 7, 2024, doi: 10.1002/aisy.202400429.
- [23] M. M. Amin, "AI and Chat GPT in language teaching: Enhancing EFL classroom support and transforming assessment techniques," *International Journal of Higher Education Pedagogies*, 2023, doi: 10.33422/ijhep.v4i4.554.
- [24] S. Ebadi and A. Amini, "Examining the roles of social presence and human-likeness on Iranian EFL learners' motivation using artificial intelligence technology: A case of CSIEC chatbot," *Interactive Learning Environments*, vol. 32, pp. 655–673, 2022, doi: 10.1080/10494820.2022.2096638.
- [25] J.-B. Jeon, "Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives," *Computer Assisted Language Learning*, vol. 37, pp. 1–26, 2022, doi: 10.1080/09588221.2021.2021241.
- [26] C. Song and Y. Song, "Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students," *Frontiers in Psychology*, vol. 14, 2023, doi: 10.3389/fpsyg.2023.1260843.
- [27] M. Li, Y. Wang, and X. Yang, "Can generative AI chatbots promote second language acquisition? A meta-analysis," *Journal of Computer Assisted Learning*, vol. 41, 2025, doi: 10.1111/jcal.70060.
- [28] D. R. C. Javier and B. L. Moorhouse, "Developing secondary school English language learners' productive and critical use of ChatGPT," *TESOL Journal*, 2023, doi: 10.1002/tesj.755.
- [29] 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, 2024, doi: 10.1007/s10639-024-12574-6.
- [30] N. P. O. Agustini, "Examining the role of ChatGPT as a learning tool in promoting students' English language learning autonomy relevant to Kurikulum Merdeka Belajar," *EDUKASIA: Jurnal Pendidikan dan Pembelajaran*, 2023, doi: 10.62775/edukasia.v4i2.373.
- [31] F. Baskara, "ChatGPT and Google Gemini in EFL education: A qualitative exploration of pedagogical efficacy among Indonesian sophomores," *Journal of Languages and Language Teaching*, 2025, doi: 10.33394/jollt.v13i1.9926.
- [32] F. Yuliani, "Multimodal feedback and student engagement: A linguistic analysis of digital assessment practices in Indonesian universities," *Indonesian Journal of Language and Educational Studies*, vol. 1, no. 1, pp. 40–55, 2026.
- [33] K. A. Maspul, V. Chemistry, Y. Melinda, K. F. Wisudayanti, and D. F. Rochmah, "How can ChatGPT empower Indonesian classrooms?," *Frontiers in Research Journal*, 2025, doi: 10.47134/frontiers.v1i3.423.
- [34] M. Tutton and D. L. Cohen, "Reconceptualizing the role of the university language teacher in light of generative AI," *Education Sciences*, 2025, doi: 10.3390/educsci15010056.

BIOGRAPHY OF AUTHOR



Wafiqotul Jamilah    is affiliated with Universitas PGRI Sumenep, Indonesia. Her academic concerns are education, language, and educational technology. Her research interests include AI-supported language learning, multilingual education, chatbot-mediated instruction, and the cultural and pedagogical implications of artificial intelligence in Indonesian classrooms. She can be contacted at email: milawafiqoh@gmail.com