www.ijarp.com ISSN 2456-9992 Page: 01-20

International Journal Advanced Research Publications

Review Article

Volume: 02
Issue: 06

GENERATIVE-AIFEEDBACKINSECOND-LANGUAGEWRITING: A CRITICAL INTEGRATIVE REVIEW AND THE PIVOT-ESL FRAMEWORK FOR LEARNER AGENCY AND ACADEMIC INTEGRITY

***DMPPK Dasanayake**

Assistant Lecturerin English, SriLanka Institute of Advanced Technological Education (SLIATE), Sri Lanka.

Article Received: 19 May 2026, Article Revised: 7 June 2026, Published on: 27 June 2026

***Corresponding Author: DMPPK Dasanayake**

Assistant Lecturerin English, SriLanka Institute of Advanced Technological Education (SLIATE), Sri Lanka.

Doi: https://doi-doi.org/101555/ijarp.8727

ABSTRACT

Generative artificial intelligence (GenAI) has rapidly entered second-language writing classrooms as a source of immediate explanations, language correction, idea development, and revision advice. Yet the pedagogical value of GenAI feedback remains uncertain because availability does not guarantee accuracy, learner uptake, or responsible use. This article presents a critical integrative review of recent empirical and design-based work on GenAI-supported writing and connects that evidence with established scholarship on second-language feedback, self-regulated learning, and feedback literacy. The review addresses three questions: what affordances and limitations characterize GenAI feedback in second-language writing; which learner and teacher practices influence productive engagement with such feedback; and how GenAI can be integrated without weakening learner agency or academic integrity. The synthesis identifies six recurring patterns: increased access to rapid and iterative feedback; useful supportforlanguageform,ideadevelopment,andconfidence;variableaccuracyandfrequentgenericity;thecentral importance of prompt and verification literacy; uneven effects on ownership and cognitive engagement; and the continuing need for teacher mediation, process evidence, and transparent disclosure. On the basis of these findings, the article proposes the PIVOT-ESL framework: Purpose and parameters, Interrogate and interact, Verify and validate, Own and operationalise revisions, and Trace, reflect, and disclose. The framework positions GenAI as a provisional feedback partner rather than an authority or substitute writer. The

article concludes that effective use depends less on the sophistication of the tool than on the quality of task design, learner judgment, teacher guidance, and institutional policy. Implications are offered for classroom practice, assessment, teacher development, and future research, particularly in resource-constrained multilingual contexts.

KEYWORDS: Generative artificial intelligence; second-language writing; feedback literacy; learner agency; academic integrity; ESL/EFL; ChatGPT.

1. INTRODUCTION

Academicandprofessionalparticipationincreasinglyrequireslearnerstoproduceclear,credible,and audience-sensitive writing in English. For second-language (L2) writers, this demand involves more than grammatical accuracy. Writers must generate and organise ideas, use disciplinary conventions, manage sources, anticipate readers, and revise across multiple levels of text. These processes are cognitively demanding even for experienced writers and are particularly challenging when learners have limited linguistic resources, uneven prior instruction, or restricted access to individual feedback.

Feedback is therefore central to L2 writing development. Well-designed feedback can help learners notice gaps, understand quality criteria, test alternative expressions, and make informed revisions. However, conventional feedback systems are often constrained by large classes, teacher workload, delayedresponse,andlimitedopportunitiesfordialogue.Writtencommentsmayalsofailtoproducel earning when they are too general, overly corrective, difficult to interpret, or delivered after students have emotionally disengaged from the task. The central challenge is not simply the provision of commentsbutthedevelopmentoflearners'capacitytointerpret,evaluate,andusethem(Carless&Bo ud, 2018; Hyland & Hyland, 2006).

The public availability of large language models has altered this feedback landscape. GenAI systems can respond to drafts within seconds, explain grammar, suggest organisational changes, simulate readers, generate examples, and continue a dialogue through follow-up prompts. These affordances are attractive in ESL and EFL settings where teacher time and access to specialist writing support are limited. Early interventionanddesignstudiessuggestthatAI-supportedwritingcanstrengthenengagement,confidence, and selected aspects of performance when learners use the tool as part of a guided revision process (Han et al., 2023; Mahapatra,

2024). At the same time, studies of generated feedback have reported abstract, generic, or inaccurate advice, especially when prompts lack task-specific criteria or when the model is asked to judge complex qualities such as coherence (Yoon et al., 2023).

A second concern is that GenAI can shift from supporting revision to replacing authorship. Learners mayacceptfluentsuggestionswithoutunderstandingthem,outsourceideageneration,orallowthem odel to reshape voice and argument. Because generated text is probabilistic rather than truth-tracking, fluent output may contain unsupported claims, fabricated references, or inappropriate stylistic choices. Thesame tool can therefore function as a scaffold, an efficiency aid, a source of misinformation, or a substitute writer depending on the task, the prompt, the learner's knowledge, and the surrounding assessment design (Kasneci et al., 2023; UNESCO, 2023).

Research on GenAI and writing has grown rapidly, but the evidence remains fragmented. Some studies examine improvement in writing scores, others analyse prompt behaviour or learner perceptions, andstill othersfocusonthe qualityofAI-generated comments.Short-term,tool-centredevaluationsoften underrepresent the processes by which learners judge feedback, retain ownership, and transfer learning to new tasks. In addition, much of the discussion treats academic integrity as a separate policing problem rather than a pedagogical issue connected to learner agency, task design, and transparent authorship practices.

The present article responds to this gap through a critical integrative review that brings recent GenAI-writing research into dialogue with established scholarship on L2 written feedback, feedback literacy, and self-regulated learning. It does not ask whether GenAI is universally beneficial or harmful. Instead, it examines the conditions under which AI-generated feedback can contribute to learning while preserving human judgment and responsibility.

Research Questions

1. Whatpedagogicalaffordancesandlimitationscharacterize GenAI-generatedfeedbackinsecond-language writing?
2. Whichlearner,teacher,task,andinstitutionalpracticesshapeproductiveengagementwithGen AI feedback?
3. HowcanGenAIfeedbackbeintegratedinawaythat strengthens learner agency, feedback literacy, and academic integrity?

Contribution of the Article

The article makes three contributions. First, it synthesises the emerging evidence around the interdependence of feedback quality, prompting, verification, uptake, and ownership. Second, it reframes GenAI use as a problem of human-AI co-regulation rather than automated correction. Third, it proposes the PIVOT-ESL framework as a practical sequence for classroom use and process-based assessment. The framework is designed for adaptable use in tertiary, adult, and upper-secondary ESL/EFL contexts, including multilingual and resource-constrained institutions.

2. CONCEPTUALFOUNDATIONS

Feed backin Second-Language Writing

L2writingfeedbackhastraditionallybeenprovidedbyteachers,peers,tutors,andautomatedwriting evaluation systems. Research has shown that feedback can support development when it is understandable, selective, timely, and connected to opportunities for revision. Nevertheless, feedback is notadirect transmission ofknowledge fromprovidertolearner.Acomment only becomespedagogically useful when the learner interprets it, decides whether it is credible and relevant, and applies it in a way that improves the text and future performance (Boud & Molloy, 2013).

The distinction between lower-order and higher-order concerns is especially important. Grammar, spelling, and sentence-level form are comparatively easy for automated systems to identify, althougheven such corrections may be contextually inappropriate. Argument quality, evidence, coherence, genre expectations, voice, and audience awareness require more situated judgment. L2 writers may be attracted to surface-level corrections because these are immediate and visible, while neglecting rhetorical and conceptual development. Effective pedagogy must therefore prevent the convenience of correction from narrowing writing to error removal.

Earlier automated written corrective feedback research offers a useful warning. Computer-generated feedback can increase opportunities for practice, but learners differ in their capacity to interpret explanations and revise accurately. Ranalli (2018) showed that automated feedback use depends on learner knowledge and the design of the support surrounding the tool. Stevenson and Phakiti (2014) similarly demonstrated that computer-generated feedback can affect writing quality, but outcomes are mediatedbythewaylearnersengagewithit.GenAIincreasesflexibilityandconversationality,yetitd oes not remove this basic dependence on uptake.

Feedback Literacy, Self-Regulation, and Agency

Feedback literacy refers to the understandings, capacities, and dispositions needed to appreciate feedback, make judgments, manage emotional responses, and take action (Carless & Boud, 2018). This construct shifts attention from the quantity of feedback to the learner's role in using it. In GenAI-supported writing, feedback literacy must include the ability to question an apparently authoritative response, compare it with assignment criteria, recognise uncertainty, and reject unsuitable suggestions.

Self-regulated learning provides a complementary perspective. Writers plan goals, monitor performance, select strategies, and evaluate outcomes in a cyclical process (Zimmerman, 2002).Feedback can strengthen this cycle when it helps learners identify a gap and choose an appropriate response. It can weaken the cycle when it removes decision-making, hides the reasoning behindrevisions, or encourages passive acceptance. Agency is therefore visible not merely when studentsoperate the tool, but when they retain control over purposes, decisions, and final wording.

GenAI also introduces co-regulation. The system can propose alternatives and ask questions, peers can compare interpretations, and teachers can guide standards and ethical boundaries. Productive human-AI co-regulation does not assign equal epistemic authority to the learner and the model. The model generates plausible language from patterns; the learner and teacher remain responsible for truth, relevance, disciplinary appropriateness, and authorship.

GenAI Feedback as Dialogic but Probabilistic

Unlike earlier rule-based feedback systems, GenAI can sustain a dialogue. A learner can request an explanation, ask for examples at a particular proficiency level, challenge an earlier answer, or supply a rubric and seek criterion-referenced feedback. This creates opportunities for iterative scaffolding and makes feedback available outside scheduled class time. It can also encourage learners who are reluctantto expose weak drafts to peers or teachers.

However, conversational fluency can create an illusion of reliability. Large language models do not independently verify the truth of a claim, understand a writer's intention in a human sense, or possess accountability for educational consequences. Their responses vary with prompt wording, model version, system settings, and contextual information. A pedagogically responsible approach must therefore treat AI feedback as provisional input to be evaluated rather than a final judgment to be obeyed.

3. REVIEWAPPROACH

A critical integrative review approach was adopted because the emerging field includes heterogeneousevidence:interventionstudies,design-baseddeployments,interactiondatasets,evaluations of feedback quality, learner-perception studies, and conceptual analyses. The purpose was explanatory synthesis and framework development rather than statistical effect estimation. Accordingly, the review does not claim exhaustive coverage in the manner of a registered systematic review.

The literature identification process focused on work published from the public release of ChatGPT in late 2022 to June 2026. Targeted searches of open scholarly indexes, academic search engines, publisher platforms, and backward reference lists combined terms such as generative AI, ChatGPT, L2 writing, EFL writing, ESL writing, automated feedback, revision, prompt engineering, learner agency, ownership, and academic integrity. Foundational research on L2 feedback, feedback literacy, automated writtencorrectivefeedback,andself-regulatedlearningwasincludedtopreventtheanalysisfromtreating GenAI as historically disconnected from earlier feedback technologies.

Sources were included when they directly addressed at least one of the following: GenAI-supported writing or revision; AI-generated feedback quality; learner prompting and interaction; cognitive,affective, or behavioural engagement with AI feedback; ownership and agency in AI-assisted writing; or integrity and governance relevant to writing assessment. Empirical and design-based studies involving L2/ELL learners were prioritised, while closely adjacent writing studies were included when they offered evidence on feedback engagement or ownership that was transferable to L2 contexts. Opinion pieces without a clear analytical basis and studies unrelated to writing or feedback were excluded from the focal synthesis.

The final focal evidence set comprised fifteen recent empirical, design-based, or evaluative sources, supplemented by established theoretical and policy literature. Each focal source was examined for context, design, role assigned to AI, reported affordances, reported limitations, and implications for learner or teacher action. Initial codes were compared and grouped through qualitative thematic synthesis informed by Braun and Clarke's (2006) principles. Themes were refined when they captured recurring relationships across different study designs rather than isolated positive or negative claims.

Because the study analysed publicly available literature and did not involve human participants or personal data, institutional ethical approval was not required. No empirical data, participant quotations, statistical results, or database counts were invented for the

present article.

Selected Evidence Base

Table1.Representative studies in forming the critical synthesis.

Study	Design/ context	Maincontribution	Cautionorimplication
Mahapatra (2024)	Mixed-methodsESLwritingintervention	Reportedimprovementinacademicwriting andpositivelearnerperceptionswhenChatGPTwasintegrated into instruction.	Potentialbenefitsdependonstructured use rather thanunrestrictedtextgeneration.
Hanetal.(2023),RECIPE	Platformdeploymentwith213 EFL students and 7instructors	Dialogicrevisionsupportgeneratedusefulinteractionpatterns and design insights.	Teacher-rolepromptingandlearnerreflection should be built into theplatform and task.
Hanetal.(2023),ChEDDAR	Semester-longinteractiondataset with 212 EFLlearners	LearnersusedChatGPTforvariedrevisionintentionsand showed differing satisfaction.	Interaction traces reveal processqualitythatfinaltextsalonecannotshow.
Hanetal.(2024),RECIPE4U	Annotatedstudent-ChatGPTdialogueandrevisiondataset	Promptintentions,satisfaction,andessayeditscanbeanalysedtogether.	AI-writingresearchshould connectdialogue behaviour with actualrevisiondecisions.
Study	**Design/ context**	**Maincontribution**	**Cautionorimplication**
Wooetal.(2023)	Caseanalysisoff ourEFLpromptpathways	Learnersfolloweddifferenttrial-and-errorrouteswhenprompting for a writing task.	Promptliteracyisteachableandshould not be assumed.
Yoonetal.(2023)	Evaluationoffeedbackon50 ELL essays	Feedbackoncoherenceandcohesionwasoftenabstract, generic, or inaccurate.	High-level discourse feedbackrequireshumanverificationandspecificcriteria.
Xuetal.(2024)	Mixed-methodsstudyofAIfeedback in translationrevision	Engagementinvolvedinteractingcognitive, affective,and behavioural dimensions.	Comprehensiblefeedbackdoesnotautomatically produce appropriaterevision.
Jelsonetal.(2025)	Onlineessay-writingstudywith 70 students	PatternsofChatGPTusewereassociatedwithenjoyment and perceived ownership.	HowstudentsuseAImattersmorethan simple access or non-accesscategories.
Ranalli(2018)	Automatedcorrective-feedbackstudy	Learnersvariedintheirabilitytouseautomatedexplanationsaccurately.	GenAIadoptionshould buildonlessons from earlier automatedfeedbackresearch.
Stevenson	Study of	Effectsonwritingqualityweremediatedbyle	Uptakeandrevisionbeh

&Phakiti(2014)	computer-generatedwriting feedback	arnerengagement with feedback.	aviourarecentral outcome variables.
Kohnkeeta l. (2023)	Language-teachinganalysis	Identified opportunities for personalised practice andrapidsupportalongsideaccuracyandethicsconcerns.	Language teachers requirepedagogicaland criticalAIliteracy.
Kasnecieta l. (2023)	Cross-disciplinaryeducationanalysis	LLMsmaysupportpersonalisedlearningbut posereliability, bias, and dependency risks.	Humanoversightandre designedassessment remain necessary.

4. FINDINGS OF THE INTEGRATIVE SYNTHESIS

Rapid Access and Iterative Dialogue Expand the Feedback Environment

The most consistent affordance is access. GenAI can provide feedback at any stage of drafting and can continue the interaction through follow-up questions. This reduces the delay between writing, noticing a problem, and attempting a revision. In contexts where teachers manage large classes, such immediacy can create additional feedback cycles that would otherwise be unavailable.

The dialogic form is pedagogically significant. Learners are not limited to a single written comment; they can ask why a sentence is unclear, request two alternatives, supply the marking rubric, or ask the model to respond as a particular audience. The RECIPE and related datasets illustrate that students use these conversations for multiple intentions rather than one uniform correction function (Han et al., 2023, 2024). Such dialogue can make revision more exploratory, but only when learners ask questions that preserve thinking rather than request a complete replacement text.

Access also has an affective dimension. Learners may be more willing to expose an incomplete or linguistically weak draft to a non-human system. Immediate, non-judgmental interaction can reduce the social risk associated with asking basic questions. However, emotional comfort should not be equated with pedagogical quality. A reassuring response may still be inaccurate, vague, or misaligned with the task.

GenAI Can Support Language Form, Idea Development, and Confidence

Recent studies suggest that GenAI can contribute to selected writing outcomes when it is embedded

inguidedinstruction.Learnerscanuseittoexplorevocabulary,comparesentencestructures,identify

recurring language errors, and receive explanations adapted to their proficiency. Because the model can generate examples on demand, it may help students bridge the gap between a rule and its application in a draft. Mahapatra's (2024) intervention study indicates that structured ChatGPT use can support ESL academic writing development and motivation.

GenAI may also support pre-writing and revision beyond grammar. It can simulate a skepticalreader, identify missing transitions, ask questions about evidence, or compare an introduction with a stated purpose. For learners who struggle to begin, a prompt-based dialogue can externalise planning and reduce blank-page anxiety. Nevertheless, the benefit is strongest when the model is used to elicit options, questions, or explanations rather than to produce a finished argument.

Confidence gains are plausible because learners receive repeated opportunities to test language privately before public submission. Yet confidence based on fluent AI output may be fragile if learners cannotreproducethelanguageindependently.Effectiveinstructionshouldthereforeincludetransfertasks in which students write without AI after guided practice, demonstrating that support has become learning rather than dependency.

Feedback Quality Is Variable, and Fluency Can Conceal Weak Judgment

The evidence does not support treating GenAI as a reliable assessor. Yoon et al. (2023) found that ChatGPT feedback on ELL writers' coherence and cohesion was frequently abstract and generic and sometimes inaccurate. This limitation is particularly important because novice writers may lack the disciplinary or linguistic knowledge needed to detect a confident but inappropriate recommendation.

Quality varies across feedback levels. Sentence-level suggestions may be locally plausible, while global advice about organisation, evidence, or genre can be based on incomplete understanding of the assignment. A model may also over-correct non-standard but meaningful expressions, flatten rhetorical voice, or impose dominant academic conventions without considering the writer's identity and audience. The result can be linguistically smoother but intellectually weaker writing.

Promptspecificity improvesrelevancebutdoesnot eliminateerror.Supplying thetask,rubric,target audience, and a request for evidence-based explanation can constrain the response. Even so, the output remains provisional. Verification is therefore not an optional final step; it is a core component of AI feedback literacy.

Prompt Literacy Shapes Both Opportunity and Inequality

Learners do not receive the same benefit merely because they use the same tool. Woo et al. (2023) documented different prompt-engineering pathways among EFL students completing a writing task.Some learners iteratively refined instructions, while others remained in unproductive trial-and-error cycles. Prompt literacy includes the ability to state a purpose, provide sufficient context, request a particular type of response, and ask follow-up questions that expose reasoning.

This creates a new dimension of inequality. Learners with stronger English, genre knowledge, and digitalconfidencearebetterpositionedtoformulateeffectivepromptsandevaluateresponses.Those who mostneedsupportmaybeleastabletorequestorjudgeit.UnstructuredAIusecanthereforewidenrather than reduce educational differences.

Prompting should consequently be taught as part of writing pedagogy, not as a technical trick. Students need models of prompts that ask for diagnosis, explanation, and alternatives while protecting authorship. They should also learn to recognise problematic requests, such as asking the system to write an assignment, invent sources, or imitate a personal voice.

Productive Engagement Requires Cognitive, Affective, and Behavioural Uptake

Feedback engagement is multidimensional. A learner may understand an AI comment cognitively, feel satisfied with it affectively, yet fail to make a useful revision behaviourally. Xu et al. (2024) found such complex relationships in AI-supported translation revision. This pattern is consistent with broader feedback research: comprehension does not guarantee action, and action does not guarantee learning.

The nature of revision matters. Copying a generated sentence can improve the submitted text while producing little development. By contrast, comparing alternatives, explaining a choice, and adapting language to the writer's own argument can strengthen knowledge and ownership. Final product scores alone may therefore overestimate learning when they do not distinguish student-generated revision from accepted AI text.

Ownership is also sensitive to usage patterns. Jelson et al. (2025) found that different forms of ChatGPT use were associated with differences in enjoyment and perceived ownership. The implication for L2 writing is that ethical and pedagogical evaluation should focus on how the tool participates in the process. A learner who uses AI to test a self-written paragraph is not engaging in the same activity as a learner who delegates planning and drafting.

Teacher Mediation Remains Essential

GenAI does not remove the teacher from feedback; it changes the teacher's role. Teachers need to design tasks that define appropriate AI functions, model questioning and verification, and create opportunities for students to explain revision decisions. They also need to address incorrect or culturally inappropriate AI advice and connect automated responses to disciplinary expectations.

Teacher mediation can occur before, during, and after AI use. Before use, the teacher clarifies learning outcomes, permitted assistance, and quality criteria. During use, learners may compare AI feedback in pairs, annotate questionable suggestions, or bring disagreements to class. After use, students can submit a revision memo showing which suggestions they accepted, modified, or rejected and why. This process converts AI interaction into visible evidence of judgment.

The need for mediation is especially strong in contexts where students view fluent English as inherentlyauthoritative.Inmanymultilingualsettings,learnersmayassumethatgrammaticallypolishedoutputisautomaticallymoreacademic.Teachersmustmakerhetoricalchoice,voice,evidence, and reader awareness explicit so that language improvement does not become uncritical standardisation.

Academic Integrity Is Best Addressed Through Process Transparency

AcademicintegritydiscussionsoftenfocusondetectingprohibitedAI-generatedtext.Detection-only approaches are unreliable and can create adversarial relationships. A more educationally defensible approach defines acceptableassistance, requires transparent process evidence, and assesses decisions that cannot be outsourced easily.

Integrity is closely connected to agency. When students understand the purpose of a task, retain responsibility for claims, and disclose the role of AI, tool use can be distinguished from ghost-writing. Revision histories, prompt excerpts, oral explanation, reflective commentary, and in-class writingsamples can provide richer evidence of authorship than an automated detector score.

Privacy and data governance are also integrity concerns. Students should not be required to upload sensitive personal information, unpublished research data, or identifiable peer writing to public systems. Institutions need clear guidance on approved tools, data retention, copyright, citation, and disclosure. UNESCO (2023) emphasises human-centred governance and the protection of learner rights; these principles are directly relevant to writing classrooms.

5. DISCUSSION: FROM AUTOMATED CORRECTION TO HUMAN-AICO-REGULATION

The synthesis suggests that the central educational question is not whether GenAI produces useful sentences. It often can. The more important question is whether interaction with the system develops the learner's capacity to write, judge, and revise independently. This distinction separates product improvement from learning development.

A tool-centred model treats GenAI as a faster feedback provider. A co-regulatory model treats it as one participant in a broader learning system that includes learner goals, teacher standards, peer dialogue, disciplinary knowledge, and institutional rules. In this system, AI contributes provisional suggestions, while human actors determine relevance, truth, ethics, and authorship. The learner remains the decision-maker, and the teacher remains the designer of conditions for learning.

This position avoids two unhelpful extremes. The first is technological solutionism: the belief that unlimited personalised feedback will automatically solve writing difficulties. The second is total prohibition: the belief that excluding GenAI will preserve conventional learning unchanged. Neither position reflects the realities of current access or the evidence on feedback uptake. A more defensible approach teaches students to use GenAI critically, selectively, and transparently.

The implications are particularly significant for Sri Lanka and comparable multilingual higher-education contexts. GenAI may increase access to language support where individual tutoring is scarce, butbenefitswillbeunevenwhenconnectivity,paid-modelaccess,Englishproficiency,andteacherAI literacy differ. Institutions should avoid policies that assume equal access or place all responsibility on individual lecturers. Shared guidance, professional development, and low-stakes experimentation are needed before AI-mediated assessment becomes routine.

6. THEPIVOT-ESLFRAMEWORK

The PIVOT-ESL framework translates the review findings into a five-phase cycle for using GenAI feedback while protecting learner agency. The word pivot signals that writers may move back to anearlier phase whenever feedback proves incomplete or unreliable. The framework is not tied to one product or model version; it describes educational actions that remain relevant as tools change.

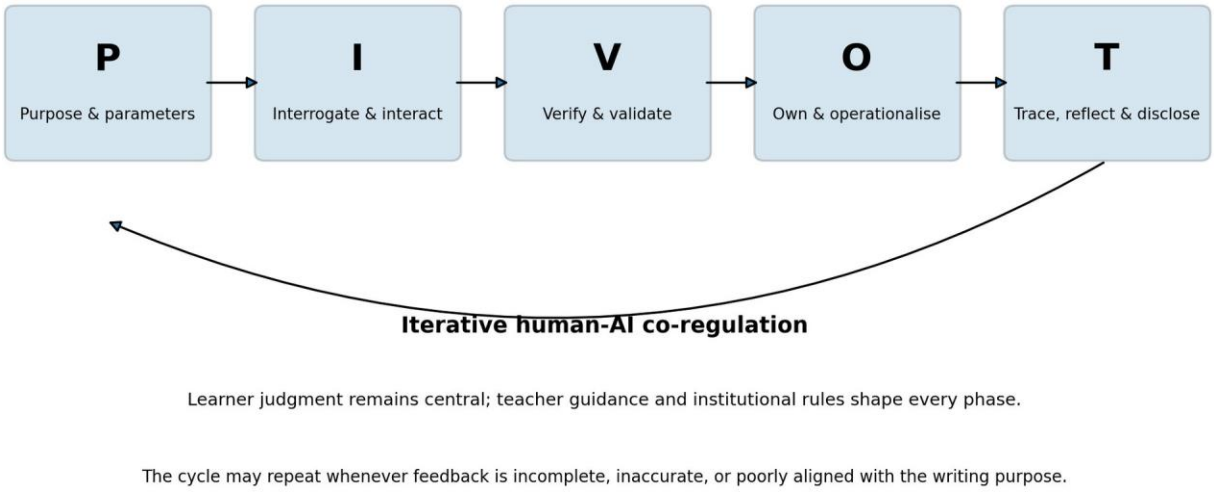

Figure1.ThePIVOT-ESLframeworkforhuman-AIco-regulatedwritingfeedback.

Table2.PIVOT-ESLphases,actions,andevidenceoflearning

Phase	**Learneraction**	**Teacher/tasksupport**	**Evidenceoflearningand integrity**
P-Purposeandparameters	Define the writing goal, audience,genre,stage ofwriting,andpermitte dAIrole.	Provide task criteria, examples ofacceptableuse,andboundaries ontextgeneration.	Alearner-writtengoalstatementandan initial draft or plan.
I-Interrogateandinteract	Ask targeted questions, requestexplanations,c omparealternatives,an dchallenge unclear advice.	Modelpromptsthatelicitdiagnosi sandreasoningratherthancomple teanswers.	Promptexcerptsshowingiterati vequestioning and task-specificinteraction.
V-Verifyand validate	Checkadviceagainstth erubric,sources, grammar references,disciplinar y knowledge, andpeer/teacherjudg ment.	Teachhallucinationawareness,s ourcechecking,andcomparisono ffeedbackproviders.	Annotationsidentifyingaccept ed,doubtful,andrejectedsugge stionswithreasons.
O - Own andoperationaliserevision s	Makeindependentdec isions,rewriteinan appropriate voice, and ensureunderstanding of every change.	Assessreasoningandtransfer,not onlysurface improvement in the final text.	A revision memo linking selectedfeedbacktochangesan dexplainingownership.
T-Trace,reflect,and	Maintainarecordofsu bstantialAI	Useprocess-basedassessmentand	Versionhistory,reflection, oral
Phase	**Learneraction**	**Teacher/tasksupport**	**Evidenceoflearningand integrity**
disclose	assistance,evaluatewh atwaslearned,and	consistentinstitutionaldisclosur eguidance.	explanation,andanAI-usestatementwhererequired.

	disclose use according to policy.		

Phase P: Purpose and Parameters

AI use should begin only after the learner understands the communicative purpose of the writing. The writer identifies the audience, genre, intended claim, assessment criteria, and current stage of work. The permitted AI role is then specified. For example, a learner may request questions about argument clarity or explanations of recurring grammatical errors but may not request a complete essay. This phase prevents the tool from defining the task on behalf of the learner.

PhaseI: Interrogate and Interact

The learner engages the system through focused, iterative prompts. Productive prompts provide context and ask the model to explain, diagnose, compare, or question. Examples include: 'Identify two places where a reader may misunderstand my reasoning and explain why'; 'Compare these two topic sentences against the rubric without rewriting them'; and 'Ask me three questions that would help me add stronger evidence.' Interrogation also means challenging the response: 'What assumption did you make?', 'Which sentence in my draft supports that judgment?', or 'Give an alternative interpretation.'

PhaseV: Verify and Validate

Learners compare AI feedback with trusted criteria and evidence. Grammar advice can be checked against references and corpus examples; content claims require credible sources; genre advice should be compared with teacher guidance and model texts. Students can mark each suggestion as accepted, uncertain,orrejectedandwriteabriefreason.Verificationturnserrordetectionintoanepistemicpractice and makes uncertainty visible.

PhaseO: Own and Operationalise Revisions

Ownership requires the writer to make and understand every final change. Learners should adapt useful advice to their own argument and voice rather than paste generated wording. A useful test is whether the student can explain the revision and reproduce the relevant principle in a new sentence or task. Teachers can strengthen ownership through revision memos, conferences, and short transfer exercises completed without AI.

PhaseT:Trace, Reflect, and Disclose

Writers preserve an appropriate record of substantial AI interaction and reflect on its value. The record need not include every minor spelling query, but it should document assistance that shaped ideas, organisation, language, or evidence. A brief disclosure statement can name the tool, purpose, and extent fuseinaccordancewithjournalorinstitutionalpolicy.Reflectionshouldaddresswhatthelearner accepted, rejected, and learned, not simply whether the tool was helpful.

7. PEDAGOGICALIMPLEMENTATION

AClassroomRevisionSequence

1. Studentsproduceanunaidedfirstdraftordetailedplansothatinitialthinkingis visible.
2. Theteacherintroducesonefeedbackfocus,suchasparagraphunity,argumentsupport,orhedging, and provides criteria.
3. Studentsuseateacher-modelledprompttoobtaindiagnosticfeedbackwithoutrequestingrewritten text.
4. Inpairs,studentscompareAIfeedbackwiththerubricandlabelsuggestionsascredible,doubtful, or irrelevant.
5. Studentsreviseselectivelyandcompleteashortmemoexplainingtwoacceptedandonerejected suggestion.
6. Afollow-upin-classtaskcheckstransferofthesameprinciplewithoutAIassistance.

Assessment Design

Assessment should reward process quality and judgment, not only linguistic polish. A portfolio can include the initial draft, selected prompts, annotated feedback, revised text, and reflection. Marks may be allocated for the accuracy of revision decisions, evidence of verification, responsiveness to criteria, and ability to explain changes. Such assessment reduces the incentive to conceal AI use because transparent use becomes part of the learning task.

Tasks can also be redesigned to include local data, personal observation, oral defence, staged submission, and comparison of sources. These features do not make AI use impossible; they make successful performance depend on situated understanding. High-stakes decisions should not rely solelyon AI-detection tools, which cannot provide dependable proof of authorship.

Teacher Professional Development

Teachers require more than technical familiarity with a chatbot. Professional development should cover prompt design, feedback literacy, model limitations, privacy, authorship, accessibility, and assessment redesign. Teachers should experience the PIVOT cycle as learners, compare outputs across prompts, and examine examples of subtle inaccuracies.

Collaborative teacher inquiry is especially valuable because acceptable use differs by discipline, proficiency, and task. Departments can develop shared prompt exemplars, disclosure language, and case-based guidance while preserving teacher judgment. Ongoing review is necessary because model capabilities and institutional policies change rapidly.

Institutional Policy and Equity

Institutional policies should distinguish assistance from substitution and explain expectations in language students can understand. A single institution-wide prohibition or permission statement is insufficient because the educational purpose of AI use varies by task. Course-level declarations should specify what is allowed, what must be disclosed, and what evidence of process may be requested.

Equity must be addressed explicitly. Some learners have access to paid models, faster devices, reliable connectivity, and extensive prior experience, while others do not. Institutions should provide equitable access to approved tools when they are required, offer non-AI alternatives, and avoid grading prompt sophistication unless it has been taught. Data protection and accessibility should be considered before adoption.

8. IMPLICATIONSFORFUTURERESEARCH

- Longitudinalstudiesshoulddistinguishimmediatetextimprovementfromdurablelearningand independent transfer.
- Comparativestudiesshouldexamineteacherfeedback,genericAIfeedback,rubric-conditionedAI feedback, and hybrid human-AI feedback under equivalent task conditions.
- Researchshouldanalyserevisionlogsandinteractiontraces,notonlyfinalscoresandself-reported satisfaction.
- Learneragency,ownership,andfeedbackliteracyshouldbetreatedasmeasurableoutcomes alongside accuracy and fluency.
- StudiesareneededinSouthAsian,African,rural,adult,vocational,andlow-resourcecontextswhere access and language hierarchies may shape use differently.

- Multilingualresearchshouldinvestigatewhenfirst-languagepromptingortranslanguagingsupports deeper reasoning and when it creates translation dependence.
- Ethicalresearchshouldexamineprivacy,bias,voicestandardisation,andtheconsequencesof requiring students to use commercial systems.
- Teacher-developmentinterventionsshouldbeevaluatedforchangesintaskdesign,feedback practice, and assessment validity rather than only attitudes toward AI.

9. LIMITATIONS

This review has four limitations. First, the evidence base is developing rapidly, and model capabilities may change faster than journal publication cycles. Second, the focal literature includes preprints and design studies because peer-reviewed L2-specific evidence remains limited; conclusions should therefore be interpreted with appropriate caution. Third, the review is integrative rather than exhaustiveanddoesnotprovideameta-analyticeffectestimateoraregisteredsystematic-searchflow.

Fourth, English-language publication bias and the concentration of studies in digitally well-resourced tertiary settings limit direct generalisation to all ESL/EFL learners.

These limitations reinforce rather than weaken the article's central recommendation: educational decisions should be guided by transparent local evidence, learner process data, and human judgment rather than broad claims about what GenAI can do.

10. CONCLUSION

GenAI has expanded the availability and form of writing feedback, but it has not removed the fundamentaleducationalproblemoffeedbackuptake.Theemergingevidenceshowsgenuinepossibilities for rapid language support, iterative dialogue, idea exploration, confidence, and additional revision opportunities. It also shows that feedback can be generic, inaccurate, rhetorically flattening, or accepted without learning. Prompt skill, subject knowledge, emotional response, teacher mediation, task design, and institutional policy all influence outcomes.

The PIVOT-ESL framework responds to this complexity by placing purpose, interrogation, verification, ownership, and transparency at the centre of AI-supported writing. Its central principle is straightforward: GenAI should contribute to the learner's decision-making process without replacing the learner as writer and accountable author. Internationally responsible adoption therefore requires morethan access to advanced tools. It requires

feedback-literate learners, critically prepared teachers, process-sensitive assessment, and policies that protect equity, privacy, and academic integrity.

For ESL/EFL education, the most promising future is neither automated writing nor a return to pre-AI practices. It is a carefully designed form of human-AI co-regulation in which technology increases opportunities for feedback while human judgment determines what counts as knowledge, quality, and responsible authorship.

DECLARATIONS

Funding: The author received no specific funding for this study.

Conflict of Interest: The author declares no conflict of interest.

Ethical Approval: Not applicable. The study analysed publicly available literature and involved no human participants or personal data.

Data Availability: No new dataset was generated. The article synthesises published and publicly accessible research.

Generative AI and AI-Assisted Technologies:

Before submission, this statement should be aligned with the target journal policy. Any use of generative AI for language support, drafting assistance, or editing should be disclosed transparently, and the author must verify every source, claim, interpretation, and final sentence.

REFERENCES

1. Bitchener, J., & Ferris, D. R. (2012). Written corrective feedback in second language acquisition and writing. Routledge. Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. Assessment &
2. Evaluation in Higher Education, 38(6), 698-712.
3. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77-101.
4. Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. Assessment & Evaluation in Higher Education, 43(8), 1315-1325.
5. Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228-239.
6. Ferris, D. R. (2003). Response to student writing: Implications for second language students. Lawrence Erlbaum Associates.

7. Han,J.,Yoo,H., Kim,Y.,Myung, J.,Kim,M.,Lim, H.,Kim,J., Lee,T.Y.,Hong, H.,Ahn,S.-Y., &Oh,A. (2023).
8. RECIPE:HowtointegrateChatGPTintoEFLwritingeducation.arXiv.https://doi.org/10.48550/arXiv.2305.11583
9. Han,J.,Yoo,H.,Myung,J.,Kim,M.,Lee,T.Y.,Ahn,S.-Y.,&Oh,A.(2023).ChEDDAR:Student-ChatGPTdialogue in EFL writing education. arXiv. https://doi.org/10.48550/arXiv.2309.13243
10. Han,J.,Yoo,H.,Myung,J.,Kim,M.,Lee,T.Y.,Ahn,S.-Y.,&Oh,A.(2024).RECIPE4U:Student-ChatGPT
11. interactiondatasetinEFLwritingeducation.arXiv.https://doi.org/10.48550/arXiv.2403.08272
12. Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. Language Teaching, 39(2), 83-101. Jelson,A.,Manesh,D.,Jang,A.,Dunlap,D.,&Lee,S.W.(2025).Anempiricalstudytounderstandhowstudentsuse
13. ChatGPTforwritingessays.arXiv.https://doi.org/10.48550/arXiv.2501.10551
14. Kasneci,E.,Seßler,K.,Küchemann,S.,Bannert,M.,Dementieva,D.,Fischer,F.,Gasser,U.,Groh,G.,Günnemann,S., Hüllermeier,E.,Krusche,S.,Kutyniok,G.,Michaeli,T.,Nerdel,C.,Pfeffer,J.,Poquet,O.,Sailer,M.,Schmidt,A., Seidel,T.,...Kasneci,G.(2023).ChatGPTforgood?Onopportunitiesandchallengesoflargelanguagemodelsfor education. Learning and Individual Differences, 103, 102274.
15. Kohnke,L.,Moorhouse,B.L.,&Zou,D.(2023).ChatGPTforlanguageteachingandlearning.RELCJournal,54(2), 537-550.
16. Mahapatra,S.(2024).ImpactofChatGPTonESLstudents'academicwritingskills:Amixedmethodsintervention study.
17. SmartLearningEnvironments,11.
18. Nicol,D.J.,&Macfarlane-Dick,D.(2006).Formativeassessmentandself-regulatedlearning:Amodelandseven principles of good feedback practice. Studies in Higher Education, 31(2), 199-218.
19. Perkins,M.(2023).AcademicintegrityconsiderationsofAIlargelanguagemodelsinthepost-pandemicera:ChatGPT and beyond. Journal of University Teaching & Learning Practice, 20(2).
20. Ranalli,J.(2018).Automatedwrittencorrectivefeedback:Howwellcanstudentsmakeuseofit?

ComputerAssisted Language Learning, 31(7), 653-674.

21. Stevenson,M.,&Phakiti,A.(2014).Theeffectsofcomputer-generatedfeedbackonthequalityofwriting.Assessing Writing, 19, 51-65.
22. Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devilismyguardianangel:ChatGPTasacasestudyofusingchatbotsineducation.SmartLearningEnvironments, 10, 15.
23. UNESCO.(2023).GuidanceforgenerativeAIineducationandresearch. UNESCO.
24. Winstone,N.E.,Nash,R.A.,Parker,M.,&Rowntree,J.(2017).Supportinglearners'agenticengagementwithfeedback: A systematic review and a taxonomy of recipience processes. Educational Psychologist, 52(1), 17-37.
25. Woo,D.J.,Guo,K.,&Susanto,H.(2023).CasesofEFLsecondarystudents'promptengineering pathwaystocompletea writing task with ChatGPT. arXiv. https://doi.org/10.48550/arXiv.2307.05493
26. Xu,S.,Su,Y.,&Liu,K.(2024).IntegratingAIforenhancedfeedbackintranslationrevision:Amixed-methods investigation of student engagement. arXiv. https://doi.org/10.48550/arXiv.2410.08581
27. Yoon,S.-Y.,Miszoglad,E.,&Pierce,L.R.(2023).EvaluationofChatGPTfeedbackonELLwriters'coherenceand cohesion. arXiv. https://doi.org/10.48550/arXiv.2310.06505
28. Zimmerman,B.J.(2002).Becomingaself-regulatedlearner:Anoverview.TheoryIntoPractice,41(2),64-70.