Indonesian Journal of Education Research (IJoER)
Vol. 7, No. 5, October 2026, pp. 663~684
ISSN: 2722-1326, DOI: 10.37251/ijoer.v7i5.3931

Designing a Bilingual English Finance Glossary for Tourism Village Financial Managers in North Sulawesi: A Conceptual Development Study

Junior Semuel Lakat[1], Daniel Semuel Warouw[2]
[1]Economics Education Department, Klabat University, North Sulawesi, Indonesia
[2]English Education Department, Klabat University, North Sulawesi, Indonesia

Article Info

Article history:

Received Aug 27, 2026
Revised Sep 29, 2026
Accepted Oct 26, 2026
Online First Oct 28, 2026

Keywords:

English For Specific Purposes
Interactive Glossary
North Sulawesi
Public Finance
Tourism Village

ABSTRACT

Purpose of the study: This study develops a literature-informed specification for an interactive English public finance glossary for tourism village financial managers in North Sulawesi. It connects financial concepts with practical English explanations and defines how the proposed resource should be reviewed and evaluated.

Methodology: A targeted narrative review of thirty-five journal articles informed a conceptual development process adapted from design science research. The work produced task-based requirements, a bilingual entry schema, thirty draft entries, learning activities, and scripted demonstrations. Content validation, usability testing, and learning assessment are specified for future research; no participant data were collected, and the glossary has not been implemented or validated.

Main Findings: The specification organizes terminology around financial tasks and distinguishes document functions, reporting concepts, and visitor transactions. It proposes a single content structure comprising bilingual definitions, examples, related terms, usage cautions, and explanatory feedback. Scripted scenarios for budgeting, payment confirmation, and reimbursement illustrate how entries are intended to align internally. These outputs are a draft design specification only; usability, vocabulary gains, and workplace transfer remain untested.

Novelty/Originality of this study: The study specifies an approach that integrates financial concept accuracy with role-sensitive English use for an intended tourism village finance audience. It proposes separating rapid reference from recognition, recall, and contextual practice, and makes institutional scope and editorial control explicit. The novelty lies in this specification and its transparent evaluation boundaries; effectiveness is not claimed and requires empirical testing.

This is an open access article under the CC BY license

Corresponding Author:
Junior Semuel Lakat
Economics Education Department, Klabat University, Arnold Mononutu Road, Airmadidi Bawah, Airmadidi District, North Minahasa Regency, North Sulawesi 95371, Indonesia
Email: jun@unklab.ac.id

1. INTRODUCTION

Tourism village financial managers work at the intersection of community finance and visitor services. They may explain an approved price, acknowledge a payment, prepare a reimbursement request, or summarize how tourism income supports a shared facility. When these activities involve English-speaking visitors or partners, a financial term must communicate the correct institutional meaning as well as the correct language form. A receipt, an invoice, and a budget are different documents; misunderstanding their functions can produce

an inaccurate explanation even when the sentence is grammatically acceptable. As tourism villages receive more visitors and partners who communicate in English, the accuracy of financial explanations becomes a practical concern for trust, transparency, and day-to-day transactions. This urgency is a conceptual argument drawn from the literature reviewed below, not an empirically established finding about North Sulawesi.

Two problems are related but not identical. The first is village financial governance: whether funds are planned, authorized, recorded, disclosed, and overseen properly. The second is financial communication in English: whether a manager can explain documents, amounts, and financial decisions accurately to English-speaking interlocutors. Strong governance does not guarantee clear English communication, and clear English cannot compensate for weak governance. This article addresses only the second problem. Responsibility for financial authorization, disclosure, and oversight remains with the relevant institutions.

The regional literature concerns the first problem. Kantale, Poputra, and Wokas examined accountability in village fund management for mangrove tourism development in Darunu, North Minahasa, and described the relevance of transparent financial processes to community empowerment [1]. Salindeho, Gazali, and Pesak report that financial transparency and human resource quality were significant predictors of tourism village development in Tumaluntung and Sawangan, whereas the positive coefficient for digitalization was not statistically significant [2]. In contrast, Kalangie, Sampe, and Kumayas describe limitations in public access to information about village-owned enterprises in West Langowan, Minahasa [3]. These studies concern different settings and cannot be combined into a province-wide estimate. They show that tourism development and village financial accountability intersect locally and that digital tools should support capable people and understandable procedures. They do not establish the English proficiency of financial managers, nor do they demonstrate demand for a digital glossary. A national analysis of village fund reporting data for 2018–2020 likewise examined accountability within reporting systems and called for more transparent and accountable fund management [4]; like the regional studies, it concerns governance rather than English communication.

Tourism village financial managers (for example, treasurers and administrators of village-owned enterprises or tourism awareness groups) are selected for three reasons. First, they handle documents and decisions on which visitors and partners rely, such as price quotations, payment confirmations, and income reports. Second, their language tasks differ from those of guides. A guide may describe an attraction, whereas a treasurer may need to distinguish a quoted price from a confirmed payment or explain why supporting evidence is required. Third, the literature on English in tourism villages has focused mainly on visitor interaction and guiding, leaving this role less examined [5], [6]. The glossary is designed to support four communication scenarios: (a) quoting and explaining prices and fees; (b) acknowledging payments and issuing receipts or invoices; (c) preparing and explaining reimbursement requests and supporting documents; and (d) summarizing how tourism income supports shared village facilities. These scenarios are derived from the cited literature and the authors' analysis of the role, not from a field-based needs analysis. Research relevant to this work falls into three strands, summarized in Table 1.

Table 1. Comparison of Related Research and Its Implications for the Present Design

Strand	Representative study	Main focus	Implication for this design	Limitation for the present purpose
ESP and bilingual materials for tourism communities	Manara and Yanti [5]; Wibowo, Nurharjanti, and Vijaya [6]	Language needs of village personnel and guides; rationale for role-based bilingual materials	Supports contextualized, task-based, bilingual content	Focus on guiding and general visitor interaction, not financial management
Needs analysis as the first stage of materials development	Suprina, Rahayu, and Sembiring [7]	Community English needs in Cibuntu; needs analysis preceding materials development	Supports a staged approach and a narrowly defined contribution	Does not address financial terminology or glossary structure
Technology-assisted vocabulary learning	Yu and Trainin [8]; Okumuş Dağdeler [9]	Overall positive effects, with variation by instructional condition and technology type; growth of mobile-assisted vocabulary research	Makes digital vocabulary support plausible	Evidence concerns language learners, not village finance personnel; effect sizes cannot be transferred
Multimedia glossing	Mahdi, Mohsen, and Almanea [10]	Outcomes vary by learner proficiency, test format, and gloss presentation	Argues for restrained media use and separate assessment of recognition and recall	Concerns general learners and general vocabulary

Taken together, these strands support contextualized bilingual materials, a staged development process, and cautious use of technology. For abstract financial concepts, an accurate definition and a contextual example may be more useful than a decorative image, which is consistent with the cognitive load principle that instructional design should limit processing irrelevant to learning [11]. Optional pronunciation audio should serve a clear communication purpose, and feedback should explain the financial distinction involved in an answer.

The comparison reveals three gaps within the selected evidence base. (1) Content gap: ESP studies in tourism villages address guiding and general community needs, while the village financial management studies address governance without designing English vocabulary resources. (2) Design gap: technology-assisted vocabulary research offers general principles, but no integrated design links financial terminology, bilingual conceptual support, tourism village workflows, and transparent evaluation for this audience. (3) Evidence gap: no field-based needs analysis of tourism village financial managers in North Sulawesi is available in the reviewed literature. We do not claim that no comparable product exists anywhere; the gaps are defined relative to the evidence reviewed here.

In response to these gaps, this article presents a conceptual development study. It does not report interviews, implementation, expert ratings, learner scores, or field outcomes, and it does not claim that user needs have been empirically established. North Sulawesi is the intended application context, not a location in which an intervention has already been tested. The article addresses three research questions:

RQ1. Which financial tasks and lexical distinctions should organize a starter glossary for tourism village financial managers?

RQ2. How can bilingual entries, search, feedback, and repeated practice be specified as one coherent learning resource?

RQ3. What evidence would be necessary to assess content accuracy, usability, vocabulary learning, and workplace relevance in a subsequent study?

The novelty of this work lies in integrating, within one reviewable specification, the financial communication tasks of tourism village managers, bilingual conceptual support, and a testable evaluation plan. Its pedagogical contribution has three parts. First, it links conceptual understanding of financial documents and decisions (for example, why a receipt, an invoice, and a budget serve different functions) to specialized vocabulary. Second, it embeds each term in application, using usage examples drawn from financial transactions such as quoting, confirming payment, and requesting reimbursement. Third, it provides feedback that explains the financial distinction behind a correct or incorrect answer. The value of the article lies in making the proposed artifact concrete and its claims testable before software deployment or empirical evaluation.

2. RESEARCH METHOD

2.1. Study Design and Boundaries

This study is a conceptual development study informed by a targeted literature review. The research object is a glossary specification, represented in the manuscript through requirements, a record schema, thirty draft lexical entries, sample feedback, and an evaluation plan. The artifact is a paper prototype: its content and interaction logic are specified, but an operational application has not been implemented or tested. The reported outcomes are therefore design outputs and analytical demonstrations, not evidence that learners have improved or that village financial practices have changed.

The study is bounded in four ways. First, it addresses financial communication in English, not village financial governance. Second, the intended users are people who handle or explain tourism-related finances within a village organization, including treasurers, administrative staff, and managers of relevant enterprise activities. This is a proposed audience definition, not a measured participant profile. Their English proficiency, accounting knowledge, device ownership, accessibility needs, and connectivity are unknown, so the starter design uses plain English with Indonesian support and leaves reading difficulty, audio needs, and terminological scope to a later needs analysis. Third, the glossary covers educational explanations of village-level financial management terms and closely related tourism enterprise transactions. It does not determine the legality of a payment, replace an accounting system, or provide an authoritative translation of every Indonesian legal term. Fourth, because village organizations differ in approval rules and reporting formats, each entry carries an institutional context field that separates terms of village government administration, village-owned enterprises, and ordinary visitor-facing transactions.

2.2. Research Procedure

The procedure adapts the design science research sequence of Peffers and colleagues [12], which comprises problem identification, solution objectives, design and development, demonstration, evaluation, and

communication. It also follows the emphasis of Hevner and colleagues on relevance and rigor in artifact-oriented research [13]. In this study, relevance concerns whether the glossary addresses identifiable financial communication tasks, and rigor concerns explicit definitions, traceable requirements, and a clear separation between published evidence and new design choices. The sequence is applied to a bounded conceptual task and is not presented as a completed design science cycle. Table 2 links each stage to the activities actually carried out, their inputs and outputs, the criterion used to treat the stage as complete, and the research question addressed.

Table 2. Research Stages, Activities Carried Out, Outputs, Completion Criteria, and Research Questions

Stage	Activities carried out	Input	Output	Completion criterion	RQ
Problem identification	Targeted literature search and selection; comparison of studies; gap analysis	Journal articles on village finance, English needs in tourism, and digital vocabulary learning	Evidence set; three gaps	Each gap is stated relative to the reviewed evidence, and each cited claim is checked against the abstract or article text	Background to RQ1-RQ3
Solution objectives	Translation of selected findings into design requirements	Evidence set	Requirement list for content, bilingual support, interaction, and evaluation	Each requirement traces to a cited finding or an explicitly stated design decision	RQ1, RQ2
Design and development	Task-to-term synthesis and entry drafting; schema design	Requirements; task scenarios	Task-based framework; record schema; thirty draft entries; interaction rules	Every entry is complete on all schema fields and meets the selection principles	RQ1, RQ2
Demonstration	Author-conducted scenario review of three scripted scenarios	Draft entries; interaction rules; scenarios	Scenario walk-throughs; list of design assumptions and failure points	Each scenario is traced through definition, context, practice question, and feedback with no unresolved inconsistency, or each inconsistency is recorded and resolved	RQ2
Evaluation	Not conducted; specified only	Draft entries; prototype	Evaluation protocol	The protocol states the evidence needed for accuracy, usability, vocabulary learning, and workplace relevance	RQ3
Communication	Preparation of this article	All outputs above	Manuscript	Not applicable	RQ1-RQ3

2.3. Literature Search and Selection

Publications were identified through targeted web searches and examination of publisher or journal pages, with institutional repositories used to resolve accessible article information when necessary. The initial searches were conducted during the preparation of the original manuscript, and the principal evidence covers publications from 2022 to 2026, supplemented by methodological and learning-science sources published from 2016 onward. A supplementary search on bilingual glossary development, ESP for accounting and finance, and digital vocabulary resources through web search and journal or publisher pages. It added eleven peer-reviewed

journal articles published between 2017 and 2024, selected for direct relevance to village fund accountability, vocabulary-learning and feedback evidence, content validation, usability evaluation, and cognitive design principles.

English expressions combined terms such as tourism village, English needs analysis, village fund accountability, mobile-assisted vocabulary learning, multimedia glosses, design science research, content validity, bilingual glossary, e-glossary, English for Specific Purposes, accounting, and finance vocabulary. Indonesian expressions included desa wisata, pengelolaan keuangan desa, transparansi, and BUMDes. Example combinations pair a topic term with a setting term, such as tourism village with English needs analysis, desa wisata with pengelolaan keuangan desa, mobile-assisted vocabulary learning with multimedia glosses, and bilingual glossary with English for Specific Purposes.

Articles were included when they met the following criteria: (a) published in a peer-reviewed journal; (b) directly relevant to at least one of six areas, namely village financial governance, English needs in tourism settings, bilingual glossary or dictionary development, ESP for accounting and finance, digital vocabulary learning, or development and evaluation methods; (c) identifiable authors and publication details, with a DOI or stable journal page; and (d) an accessible abstract or full text. Articles were excluded when they were not peer-reviewed, when they had no connection to village, tourism, financial, or vocabulary-support contexts, or when no abstract or text was accessible for verification. Relevance and the quality of the evidence, not recency, guided selection. Quality was judged by whether the study reported an identifiable setting, method, and finding, and whether the finding was relevant to a specific design decision.

Thirty-five articles were retained as a purposeful evidence set for artifact development. Twenty were published from 2022 onward, which reflects the focus on current tourism village and digital vocabulary research; thirteen were published between 2016 and 2021; recency is reported descriptively and is not treated as evidence of quality. Two earlier sources were retained because they provide established methods for artifact development that are not superseded by newer topic-specific studies: Peffers and colleagues [12] and Hevner and colleagues [13]. The 2016–2021 group includes Yusoff [14] for content validation, Rohani and Suyono [15] for the development of an English-Indonesian ESP e-glossary, and Nugroho [16] for needs analysis of English materials in accounting, together with the supplementary sources on vocabulary learning, feedback, practice testing, cognitive load, content validity, and usability measurement.

This is a targeted narrative review, not a systematic review. No database-wide coverage, complete search yield, duplicate-removal count, or exhaustive screening flow is claimed, and the thirty-five references are not a representative sample of all research in the area. Claims were checked against the accessible abstract or article text. Where only abstract-level evidence was available, the article reports the broad finding and avoids interpreting instruments or analyses that were not examined. Publication metadata were verified on journal or publisher records. The samples and outcomes of the cited studies are not represented as observations made in this study.

2.4. Requirements Synthesis and Lexical Development

Development moved from financial task to communicative act, and from communicative act to terminology and learning activity. A payment task, for example, involves requesting or confirming money, identifying its purpose, and recording evidence, which leads to invoice, receipt, advance payment, and supporting document. A reporting task involves identifying the reporting period and explaining differences between planned and actual values, which leads to budget, budget realization, reporting period, and variance. This reasoning produced the starter set. It is not a frequency analysis of an authentic local corpus. Studies of accounting students report difficulty with financial terminology in reports, auditing, and budgeting [17], which is consistent with prioritizing terms attached to financial documents, although those findings concern university students and are not treated as evidence about village managers.

Terms were selected and prioritized using four initial development principles. These are design principles applied by the authors, not the result of an empirical assessment: (1) task relevance, meaning the term is needed in at least one of the four communication scenarios; (2) potential for misunderstanding, meaning the term is easily confused with a related term, as with receipt, invoice, and budget; (3) communication requirement, meaning an English-speaking visitor or partner would need to understand or use the term in a transaction; and (4) terminological specificity, meaning the term has a precise institutional meaning that a general dictionary equivalent would not convey. Terms meeting more principles were given priority. The principles were applied through the authors' judgment, without weighting or a formal checklist.

The thirty entries illustrate distinct functions across six working domains: planning, revenue, expenditure, cash and assets, reporting and oversight, and visitor transactions. The number is a design scope decision for a manageable first version and does not estimate the vocabulary required for professional competence. The English definitions were drafted by the authors in original plain English, and no definition was adapted from a specific published glossary or dictionary. The Indonesian equivalents are working equivalents drafted by the authors and were not verified against an official or legal translation. Each entry was connected to a

fictional tourism village example. Technical terminology was retained where necessary, and circular definitions and unexplained synonyms were avoided.

The schema separates meaning, language use, institutional scope, and learning support, so that a short translation is not made to carry several responsibilities. The English definition explains the concept, the Indonesian equivalent supports access, the example sentence shows a communicative use, the caution identifies a likely confusion, and the review fields document editorial control. The pairing of English terms with Indonesian descriptions and example sentences is consistent with the structure of an existing English-Indonesian ESP e-glossary for polytechnic students [15]; the present schema differs by adding institutional context, caution, and review fields for the village financial setting. Table 5 presents the schema and Table 6 the thirty-entry starter set.

2.5. Data Analysis

The study does not analyze participant data. The analyzed materials are the thirty-five selected articles and the draft glossary content. The analysis was a qualitative, author-conducted synthesis, carried out in the following sequence:

1. Extraction. For each article, the setting, focus, main finding, and limitation were recorded in an evidence matrix.
2. Comparison and gap identification. Studies were grouped into six strands (bilingual glossary and dictionary development, ESP for accounting and finance, ESP and bilingual materials for tourism communities, needs analysis, technology-assisted vocabulary, and multimedia glossing) and compared on content, design, and evidence.
3. Requirement derivation. Each finding that bears on design was translated into a requirement. Published findings are used as design reasoning only, and effect sizes are not transferred to the target group.
4. Task-to-term mapping. The four communication scenarios were broken into communicative acts and then into candidate terms.
5. Term selection. Candidate terms were screened against the four selection principles, and thirty were retained.
6. Entry construction. Retained terms were drafted into schema records, and cautions were written for terms with likely confusions.
7. Scenario review. Draft entries and interaction rules were examined through scripted scenarios.
8. Reporting. Results are reported as design outputs. Observations from the scenario review are interpreted as design assumptions and failure points, not as evidence of learning or usability.

The synthesis was conducted by the authors without independent coding or an external check, so it reflects the authors' judgment and has not been tested for agreement between independent coders.

2.6. Analytical Demonstration

The paper prototype was examined through three scripted scenarios: a budget discussion, a payment confirmation, and a request for financial evidence. For each scenario, the authors wrote out the expected user goal, the entries and navigation steps involved, the expected answer, and possible misunderstandings. The review examined four aspects: (1) whether the definition accurately conveys the institutional meaning; (2) whether the usage context matches the scenario; (3) whether the practice question tests the distinction that the definition explains; and (4) whether the feedback explains the financial distinction behind a correct or incorrect answer.

A scenario was treated as consistent when the definition, context, practice question, and feedback addressed the same distinction without contradiction, and when a plausible misunderstanding had a corresponding caution or feedback message. Any inconsistency found was treated as a design finding to be resolved in the entry or rule concerned. The review was a conceptual assessment conducted by the authors only. No user executed the tasks, no external expert reviewed the content, and no completion times, usability scores, or success rates were collected. The purpose was to expose design assumptions and failure points before implementation, and the review is not equivalent to participant testing.

2.7. Planned Validation and Evaluation

The following procedures are specified for a subsequent study and were not carried out here. A local needs analysis should combine role descriptions, examples of permitted workplace documents, observations of actual communication, and learner accounts of difficult terms. Pang, Abd Majid, and Narayanan's review supports the use of multiple information sources and attention to disciplinary and contextual differences in English for Specific Purposes needs analysis [18]. Prior ESP dictionary development has likewise begun with a needs analysis that combined a student survey with interviews of lecturers and industry practitioners [19], and needs analyses of accounting students' English materials have been reported in Indonesia [16]; these precedents

concern students, so the instruments would need adaptation for village financial managers. Interviews alone would be insufficient if not checked against the texts and financial tasks users actually encounter. Access to documents would require permission and removal of personal or sensitive financial information. Evaluation of the implemented glossary should then assess content accuracy, usability, vocabulary learning, and workplace relevance.

3. RESULTS AND DISCUSSION

This section reports what the conceptual development produced and discusses it against prior studies. Because no software has been built and no user has been tested, every statement carries one of three statuses. Output means an artifact or analysis that exists in this article, such as the record schema, thirty draft entries, three model entries, and three scripted scenarios. Proposal means a design decision that is specified here but not implemented, such as offline delivery, review intervals, and search ranking. Pending means a result that can be determined only through empirical testing, such as learning gains, usability, content validity indices, and the local relevance of the terms. Prior studies are used for comparison only, and their findings and effect sizes are not treated as results of this study. Table 3 summarizes the register. The three research questions are answered by design findings, not by empirical findings about learners: RQ1 by Sections 3.1 to 3.3, RQ2 by Sections 3.4 to 3.8, and RQ3 by Sections 3.9 to 3.11.

Table 3. Status Register of the Principal Elements

Element	Status	Reported in	What would change its status
Evidence-informed requirements	Output (interpretive mapping)	3.1	Local needs analysis
Record schema	Output	3.2	Test of editorial use
Thirty draft entries	Output (not yet reviewed)	3.3	Content validation [14]
Three expanded model entries	Output	3.4	Learner comprehension testing
Search and navigation rules	Proposal	3.5	Functional prototype testing
Activities and review schedule	Proposal	3.6	Learning study
Three scripted scenarios	Output (author review only)	3.7	Participant testing
Offline and accessibility requirements	Proposal	3.8	Acceptance tests
Validation and evaluation protocol	Proposal	3.9, 3.10	Execution of the protocol
Anticipated failure cases	Output (anticipatory)	3.11	Observed error data
Learning gains, usability, content validity, workplace relevance	Pending	none	

3.1. Evidence Translated Into Design Requirements

Table 4 maps the principal requirements to selected evidence and to proposed verification. The mapping is an interpretive output of this study: the publications do not directly prescribe the exact glossary design. Each proposed feature is a development choice made for the intended task (Proposal), and the last column marks evidence that remains Pending. In particular, offline access is a precautionary design assumption that must be checked through local needs analysis. It is not a reported finding that all tourism village managers in North Sulawesi have unreliable internet access.

Table 4. Evidence-Informed Requirements and Proposed Verification

Requirement	Evidence or Rationale	Proposed Feature (Proposal)	Evidence Still Needed (Pending)
Explain financial purpose	Regional governance studies [1]	Definition with document function and scope	Local finance reviewer approval
Support transparent explanation	Organizational capacity study [2]	Examples of reporting and supporting evidence	Actual user communication tasks

Make information understandable	Public access limitations [3]	Plain explanations with institutional cautions	Comprehension testing
Relate language to roles	Tourism English needs study [4]	Browse by work task and document	Local role and text analysis
Separate learning outcomes	Glossing synthesis [10]	Recognition, recall, and use activities	Independent assessment results
Support continued practice	Spacing meta-analysis [20]	Optional scheduled review	Retention and adherence data
Evaluate real interaction	Usability mapping study [17]	Observed tasks and error records	Functional prototype testing
Pair English terms with Indonesian explanations and examples	English–Indonesian ESP e-glossary development [15]	Bilingual entry with working equivalent and example sentence	Local review of the equivalents
Prioritize confusable financial terms	Difficulty with financial terminology in reports, auditing, and budgeting among accounting students [21]	Priority for terms whose confusion changes an explanation	Corpus of authorized documents and observed communication
Base term selection on needs analysis	Needs analyses of accounting students [16] and of air transport students for a dictionary [19]	Local needs analysis before the term set is expanded	Role, text, and task data from partner organizations

3.2. Glossary Architecture and Editorial Schema

The proposed architecture has two layers. A reference layer provides rapid access to meanings and examples, while a practice layer provides retrieval tasks and contextual decisions. Both layers use the same approved record. This reduces the risk that a quiz teaches one meaning while the reference entry displays another. A stable identifier connects the headword, search aliases, examples, audio, and assessment item. If an entry changes, associated learning activities must be reviewed before the revised record is released.

Table 5 specifies the record fields. Some fields are learner-facing, while others support editorial work. A small initial interface need not display all editorial fields at once, but the underlying content record should retain them. Separating fields also makes it possible to produce a printed reference sheet without exporting participant information or internal comments. The schema is a specification for a future content system and does not imply that a database already exists.

Table 5. Proposed Record Schema

Field	Purpose	Illustration or Rule
Entry identifier	Maintain stable links	PF-001 remains stable after a wording revision
English headword	Identify the concept	Budget; one concept per primary record
Indonesian equivalent	Support bilingual access	Anggaran; label as a working equivalent
Plain English definition	Explain the core meaning	Planned income and spending for a stated period
Institutional context	Limit misleading transfer	Village administration, enterprise, or general use
Example sentence	Show usable English	The budget includes a maintenance allocation.
Document or task tag	Connect to work	Planning; financial report; visitor payment
Related terms	Clarify contrasts	Expenditure; budget realization; variance
Usage caution	Identify a likely confusion	A budget is a plan, not proof that money was spent.
Learning prompt	Require an active response	Select the document that records actual payment.
Audio and transcript	Support pronunciation	Optional, reviewed recording with visible text
Review and version fields	Support editorial control	Reviewer role, review date, source note, version

The English definition should normally fit within one or two short sentences. Brevity is a target rather than an absolute word limit, because removing a necessary condition can make a financial explanation wrong. Definitions should use common verbs such as plan, receive, pay, record, compare, and check whenever these accurately describe the concept. A specialist term introduced inside a definition should link to its own entry or be

explained immediately. The learner should not have to move through several unknown words to understand the first one.

Examples should state who acts, what happens, and which document or period is involved. The sentence 'The cashier issued a receipt after receiving the payment' is more useful than 'A receipt is important.' The first sentence models both the word and its relationship to an event. Examples should avoid unrealistically complex grammar during early learning, but grammatical simplification should not erase dates, amounts, conditions, or authorization boundaries. Financial communication often becomes misleading precisely when these details disappear.

3.3. Starter Vocabulary and Institutional Distinctions

The second development output is the thirty-entry starter set in Table 6. The table provides compact draft definitions and example uses, allowing the lexical scope to be reviewed before implementation. It is not an official accounting dictionary or an approved local training syllabus. The examples are original fictional teaching sentences. Institutional terms and translations require local review, especially where an organization uses a prescribed Indonesian reporting vocabulary.

Table 6. Thirty Draft Entries For the Starter Glossary

English term and working Indonesian equivalent	Plain English definition	Fictional example sentence
Budget anggaran	A plan of income and spending for a stated period.	The annual budget includes funds for pathway maintenance.
Budget allocation - alokasi anggaran	An amount in a budget assigned to a purpose.	The committee discussed the allocation for waste collection.
Budget realization - realisasi anggaran	Actual budget execution reported against the planned amounts.	The report compares budget realization with the approved plan.
Revenue - pendapatan	Income recognized under the applicable reporting rules.	The report identifies tourism service revenue for the period.
Expenditure - belanja or pengeluaran	Spending for an organizational purpose; the formal category depends on the reporting framework.	The expenditure summary shows payments for maintenance work.
Grant - hibah	Funding provided for an agreed purpose under stated conditions.	The grant agreement requires a final activity report.
Own source revenue - pendapatan asli	Revenue generated from an entity's own sources rather than external transfers; scope must be specified.	The report identifies the village government's own source revenue.
Procurement - pengadaan	The process of obtaining goods or services for an organization.	The procurement file contains the approved specifications.
Purchase order - surat pesanan	A document specifying goods or services an organization orders.	The purchase order states the quantity of cleaning supplies.
Invoice - tagihan or faktur	A document stating an amount requested or due for goods or services.	The supplier sent an invoice for the repair work.
Receipt - kuitansi or tanda terima pembayaran	A document acknowledging that a payment has been received.	The cashier gave the visitor a receipt after payment.
Supporting document - dokumen pendukung	Evidence used to substantiate a transaction or report.	Please attach the supporting document to the request.
Cash flow - arus kas	The movement of cash into and out of an organization over a period.	The summary shows cash flow for the month.
Cash balance - saldo kas	The amount of cash recorded or held at a specified time; specify which balance is meant.	The treasurer checked the recorded cash balance at closing.
Bank reconciliation - rekonsiliasi bank	A comparison of accounting records and a bank statement to explain differences.	The bank reconciliation explains the unmatched payment.
Asset - aset	A resource controlled by an entity under the applicable accounting framework.	The equipment register records the new asset.

Liability - kewajiban	An obligation to transfer resources under the applicable accounting framework.	The report identifies the unpaid supplier amount as a liability.
Surplus - surplus	An excess of income or revenue over expenses or expenditure on a stated basis and for a stated period.	The manager explained the reported surplus for the period.
Deficit - defisit	An excess of expenses or expenditure over income or revenue on a stated basis and for a stated period.	The committee discussed how to address the reported deficit.
Internal control - pengendalian internal	Organizational procedures designed to manage risks and support reliable operations and reporting.	Separate payment approval is part of internal control.
Accountability - akuntabilitas	The obligation to explain and justify actions and the use of entrusted resources.	The report supports accountability to the community.
Transparency - transparansi	The availability and clarity of relevant information for those entitled to access it.	A clear public summary can support transparency.
Audit - audit	A structured examination of records or activities against stated criteria.	The audit request asks for the transaction evidence.
Financial report - laporan keuangan	A document presenting financial information for an entity and period; its form depends on the framework.	The financial report states the reporting period.
Reporting period - periode pelaporan	The interval covered by a financial report.	This reporting period runs from January to March.
Entrance fee - biaya masuk	A charge for admission to an attraction or facility under its approved arrangement.	The entrance fee covers admission to the demonstration site.
Refund - pengembalian pembayaran	Money returned to a payer under the applicable policy or agreement.	The visitor requested a refund after the cancellation.
Reimbursement - penggantian biaya	Repayment of an eligible expense previously paid by another person.	The staff member requested reimbursement for approved transport costs.
Advance payment - pembayaran di muka	Money paid before the related goods or services are fully delivered.	The invoice records the advance payment already received.
Variance - selisih	The difference between compared values, such as planned and actual amounts; the calculation basis must be stated.	The note explains the variance between planned and actual spending.

Several entries deliberately contain a context condition. Revenue, expenditure, asset, and liability have meanings that depend partly on the accounting framework. An introductory glossary should support understanding without pretending that one short definition captures every recognition rule. The phrase 'under the applicable reporting rules' signals that the entry is educational. It also gives the instructor a reason to connect the word to the documents and procedures actually used by the organization.

Own source revenue requires special care. In village administration, a related Indonesian expression may refer to a formal category of village income. In an enterprise, sales revenue belongs to a different reporting context. The English expression should not be used to merge these categories. A future version should create distinct records or explicit context notes if users need both meanings. The starter entry therefore specifies the entity rather than equating every locally generated tourism payment with one legal budget classification.

Surplus and deficit also require a stated basis. A positive cash balance does not necessarily establish a surplus, because the timing of receipts, payments, and obligations may differ from the recognition of income and expenses. The glossary should model cautious explanations such as 'This is the cash balance at the end of the month' when that is the known fact. It should avoid encouraging users to call the same figure a profit, surplus, available budget, and cash balance interchangeably.

Transparency is framed as availability and clarity for those entitled to access information. This allows the learning material to discuss understandable public summaries without suggesting that every individual payment detail must be disclosed to every visitor. The entry can connect transparency to document access, explanation, and appropriate disclosure, while accountability connects to justification and responsibility. The

distinction helps learners explain why a financial report may be public in summarized form even when personal identifiers or protected transaction details are withheld. The starter set omits tax rates, statutory deadlines, procurement thresholds, and actual village tariffs. A later version adding these values must identify their jurisdiction, effective date, and authoritative source, with a process for withdrawing outdated content.

3.4. Model Entries and Bilingual Explanation

Three expanded records illustrate how definitions, translations, contextual sentences, comparisons, and feedback function together. They specify the content logic that a later application should preserve. For receipt, the English definition is 'A document acknowledging that a payment has been received.' The working Indonesian equivalent is kuitansi or tanda terima pembayaran. The example sentence is 'The cashier issued a receipt after receiving IDR 150,000 for the group booking.' The task tag is confirm a payment, and the related term is invoice. The usage caution states that a receipt acknowledges payment, whereas an invoice normally requests or records an amount due. A document's actual contents must still be checked, because some organizations use combined documents.

The receipt learning prompt presents a short sequence: a group receives an amount due, pays the amount, and asks for evidence of the payment. The learner selects the appropriate document term. If the learner selects receipt, feedback explains that the payment has already occurred and that the requested document acknowledges it. If the learner selects invoice, feedback explains the different function and directs the learner to compare the two entries. This feedback addresses the underlying confusion, so the learner receives more than a correct-or-incorrect signal. For budget allocation, the definition is 'An amount in a budget assigned to a purpose.' The Indonesian equivalent is alokasi anggaran. The example states, 'The proposed budget allocates IDR 4,000,000 to pathway maintenance.' The caution explains that an allocation is not proof of expenditure and does not by itself establish cash availability. The learner is asked whether the sentence shows a plan, an actual payment, or a bank balance. The expected answer is a plan. The feedback identifies the words proposed budget and allocates as the clues.

For bank reconciliation, the definition is 'A comparison of accounting records and a bank statement to explain differences.' The working equivalent is rekonsiliasi bank. The example states, 'The treasurer is checking a payment recorded in the ledger but not yet shown on the bank statement.' The caution explains that a difference should be investigated rather than immediately classified as fraud or corrected by deleting a record. The entry is tagged to cash and bank review, with related terms cash balance and supporting document. This example introduces a careful explanation without providing a complete accounting procedure.

The bank reconciliation activity asks the learner to complete the sentence 'We are comparing the ledger with the bank statement to explain the difference.' The activity checks vocabulary use, not mastery of reconciliation accounting. A follow-up question asks what additional evidence would help clarify the difference. A model response refers to the payment date and supporting document. The resource should state that actual corrections require the organization's authorized procedure, because an educational answer does not authorize changing financial records.

Bilingual explanation should preserve the conceptual difference between terms. If an Indonesian explanation for receipt says only dokumen transaksi, it may be too broad to distinguish the term from invoice. A stronger explanation refers to acknowledgment of a payment that has been received. The same principle applies to refund and reimbursement. Both involve money moving back to someone, but their reasons differ: a refund returns a payer's money under a policy or agreement, while reimbursement repays an eligible expense previously incurred by another person. Meta-analytic evidence that first-language glosses produced greater vocabulary learning than second-language glosses [22] supports retaining the Indonesian support, although that evidence concerns general learners rather than financial terminology.

Grammar support is embedded in useful sentence patterns. Examples include 'The invoice shows ...,' 'The receipt confirms ...,' 'The report covers ...,' and 'The amount is allocated for ... .' These patterns offer a manageable way to communicate a financial fact while leaving room for an amount, purpose, or period. The glossary should encourage complete, precise sentences rather than isolated translated words. An instructor can later expand the patterns into email requests, brief explanations, and report notes after checking the tasks users actually perform.

3.5. Search and Navigation Specification

The reference layer should support search in English and Indonesian. A search for kuitansi should return receipt, while a search for tagihan should offer invoice with a context note. The system should show the primary headword and a one-line definition before the user opens the full record. This allows the learner to inspect the likely meaning without committing to the first result. Search should not silently treat all semantically related expressions as exact equivalents.

The proposed matching order is exact headword, approved alias, normalized spelling variant, and broader related term. Exact results should appear before suggestions. For a misspelling, the interface should

explain that it is offering a possible match rather than rewriting the user's intention without notice. A broad query such as payment should produce several relevant entries with distinct descriptions, because payment can relate to an invoice, a receipt, an advance, or a refund. Ranking rules require testing against realistic user queries after implementation. When no match exists, the proposed screen displays a clear message and offers task categories or related terms. It should not generate an unreviewed definition automatically. A request to suggest a new term can record the headword and a non-sensitive example, subject to later review. The interface should advise users not to upload actual bank statements, personal identifiers, or transaction records into a public suggestion form. A glossary needs examples of communication, but it does not need confidential financial evidence to provide its ordinary reference function.

Task browsing should use short labels that correspond to recognizable actions. The initial categories are plan spending, explain income, document a payment, check cash and bank records, prepare a report, and explain a visitor transaction. These labels can be revised during needs analysis. If users organize their work by document rather than task, a parallel document route can offer budget, invoice, receipt, and financial report. A fixed information structure should remain open to revision when it conflicts with users' actual vocabulary for their work. The entry screen should present the English term, Indonesian equivalent, and core definition first. The example sentence, related terms, pronunciation audio, and practice prompt can appear below. The usage caution should remain visible whenever it concerns a likely misunderstanding that changes the meaning of a transaction. Hiding every caution behind an optional expansion would make the interface shorter but could remove information needed for a correct explanation. Conversely, extensive editorial notes should stay outside the main learning view.

3.6. Interactive Learning and Feedback

The practice layer separates recognition, recall, and contextual use. Learners identify a meaning, retrieve a term without its complete form, or explain a financial situation in an appropriate sentence. Later evaluation should report these outcomes separately because recognition alone does not establish independent production. Research on spaced practice provides a basis for investigating repeated review rather than a single exposure. Kim and Webb report benefits of spacing in second language learning while showing that outcomes depend on methodological and instructional conditions [20]. Meta-analytic evidence on practice testing similarly reports that practice tests were more beneficial for learning than restudying [23], and a meta-analysis of repetition effects on incidental vocabulary learning [24] motivates recording how often learners meet each entry. The proposed glossary can offer voluntary review of selected terms after an initial learning session. A provisional schedule might revisit difficult entries after one day, three days, and one week. These intervals are design choices for piloting; they are not presented as the uniquely optimal schedule derived from the review.

Longer-term mobile vocabulary evidence also supports attention to sustained engagement. Zhou and Zhou's meta-analysis reports positive vocabulary outcomes for mobile applications and examines factors associated with those outcomes [25]. Dan and colleagues similarly identify intervention duration as relevant in their synthesis [26]. For this glossary, the practical implication is to evaluate how often and how long learners actually use the resource. It does not follow that a short tourism finance intervention will reproduce outcomes from longer studies in different educational settings.

Table 7 presents the proposed activity types. The glossary can begin with a small set rather than implementing every type at once. The principal design constraint is that the response and feedback must refer to a clear financial distinction. Activities that reward superficial clicking without requiring interpretation would add usage counts without showing useful learning. The proposed model avoids treating the number of completed screens as a direct measure of competence.

Table 7. Proposed Activities and Feedback Logic

Activity	Prompt format	Target outcome	Feedback focus
Meaning recognition	Choose the definition of invoice.	Identify core meaning	An invoice states an amount requested or due.
Term comparison	Select invoice or receipt after a payment.	Distinguish document functions	Payment acknowledgment differs from a request for payment.
Meaning recall	Name the return of a visitor's payment.	Retrieve refund from context	Explain why reimbursement fits a different situation.
Sentence completion	The report covers the reporting	Use a financial expression	Connect period to the interval covered.
Document interpretation	Locate the amount due in a fictional invoice.	Read a financial text	Distinguish total charge, advance, and balance due.
Brief explanation	Explain an actual spending variance.	Produce contextual language	State the comparison basis without unsupported conclusions.

Feedback should be brief enough to read during a task but complete enough to resolve the misconception. For an incorrect receipt answer, the message can say, 'A receipt confirms that payment was received. The supplier is asking for payment, so invoice fits this situation.' This wording explains both the selected word and the expected word. A meta-analysis of educational feedback reports that its impact depends substantially on the information content it conveys [27], which supports feedback that names the financial distinction. A second attempt can present a different example rather than allowing memorization of the same response position. The system should vary option order only if that does not confuse accessibility tools or increase unnecessary cognitive effort.

Audio should be optional and accompanied by the written headword and sentence. Users may study in a shared office, at a service counter, or without headphones. The design should not start audio automatically or make listening a condition for reading a definition. A recording should prioritize intelligibility and consistent pronunciation of the English term. The resource can acknowledge acceptable pronunciation variation without judging local accents as financial incompetence. A pronunciation task and a concept interpretation task should be assessed separately.

Digital reading evidence supports connecting vocabulary to short texts rather than keeping every term isolated. Zhu, Zhang, and Irwin review vocabulary learning through digital reading and identify variation associated with learner and resource characteristics [28]. In the proposed glossary, a fictional report excerpt can contain links to budget, expenditure, and variance. The links should allow a learner to inspect the meaning while retaining the text context. This is a design inference from the literature, not a tested finding about tourism village reports.

Mohsen, Mahdi, and Alkhammash also show that the effects of multimedia glosses vary with proficiency and assessment type [29]. A meta-analysis of word-focused activities such as flashcards, word lists, writing, and fill-in-the-blank tasks likewise found wide variation in learning gains across activities and lower gains on delayed than on immediate tests [30]. These findings reinforce the need to distinguish exposure from demonstrated learning. A learner who opens several glosses may be engaged, but opening does not show recall or appropriate use. The practice layer should therefore collect responses to meaningful tasks when users consent to evaluation, while the ordinary reference layer remains available without requiring an assessment session.

For open responses, the first implementation should use an instructor-reviewed rubric or a small set of transparent acceptable patterns. Automated scoring can misclassify a grammatically imperfect but financially correct explanation or accept a fluent but misleading one. The system should not label an open response as financially correct merely because it contains the target headword. A human reviewer must consider the amount, period, condition, and document relationship in the scenario. This requirement limits claims about how much of contextual learning can be evaluated automatically.

Progress indicators should describe the activity completed rather than promise mastery. Labels such as reviewed, answered correctly in practice, and scheduled for review are more defensible than certified or financially competent. Users should be able to revisit an entry even after answering correctly. An instructor can review patterns of misunderstanding without exposing a public ranking. The intention is to support practical learning and diagnosis, while avoiding an incentive to rush through questions for a score that has little relationship to work.

3.7. Worked Financial Communication Scenarios

The examples presented in this section are based on an author-conducted scenario review using fictional values to examine alignment between entries, task routes, and feedback. They are analytical examples rather than observations of participant performance. In the first scenario, a tourism committee discusses an annual maintenance plan. The proposed allocation is IDR 12,000,000, and a later fictional summary reports IDR 9,600,000 in actual spending. The learning task asks the user to distinguish the budget allocation from expenditure and describe the difference. Under the stated calculation convention, planned spending minus actual spending equals IDR 2,400,000. The figure is a spending variance in the example, not automatically a cash balance, profit, or unrestricted amount available for another purpose.

The reference path begins with plan spending and leads to budget allocation. The related terms offer expenditure and variance. A model English explanation is, 'The planned maintenance allocation was IDR 12,000,000. Actual spending was IDR 9,600,000, which was IDR 2,400,000 below the planned amount.' A second sentence can explain the reason only if the scenario provides evidence. The glossary should not encourage the user to add 'because the manager was efficient' or 'because the project was incomplete' when neither conclusion follows from the figures alone. This scenario reveals two requirements. The variance record must specify the calculation direction, because actual minus planned would produce a negative sign for the same difference. The resource must also distinguish a numerical description from an evaluation of performance. Below-budget spending can have different explanations, and the glossary cannot determine which applies.

Feedback should reward the accurate comparison and direct the learner to supporting information for any explanation of cause.

In the second scenario, a group booking costs IDR 600,000. The group has already paid an advance of IDR 200,000, leaving a balance due of IDR 400,000 under the stated fictional agreement. Before final payment, the visitor asks how much remains payable. After payment, the visitor asks for evidence that the balance was received. The relevant terms are invoice, advance payment, and receipt. The scenario makes their sequence visible rather than teaching them as disconnected translations. The model explanation before payment is, 'The total charge is IDR 600,000. We have received your advance payment of IDR 200,000, so the remaining amount due is IDR 400,000.' After payment, the explanation becomes, 'This receipt confirms that we received the remaining IDR 400,000.' The resource should preserve the currency code and amount. It should not introduce a tax, discount, or cancellation condition unless the fictional document explicitly provides one. The task evaluates reading and explaining given information rather than deciding the organization's pricing policy.

An incorrect response that calls the invoice proof of final payment receives a comparison prompt. The learner is asked to identify which event has occurred: an amount has been requested or money has been received. If a later local document combines both functions, the instructor should explain that the document's status and contents control the interpretation. The glossary should allow this qualification, because real documents do not always follow a single idealized layout. The starter scenario teaches the common distinction while leaving space for verified local practice.

In the third scenario, a staff member pays IDR 180,000 for approved transport and requests repayment. The fictional organization requires evidence of the expense and confirmation that it was eligible. The target terms are reimbursement and supporting document. The model sentence is, 'Please attach the receipt and the approval record to your reimbursement request.' This sentence teaches a financial communication function, but it does not assert that these two documents are always legally sufficient in every organization.

The feedback contrasts reimbursement with refund. A visitor seeking return of a payment after cancellation is a different situation from a staff member requesting repayment of an eligible expense. Both involve a transfer back to a person, but the reasons and evidence differ. The model entry therefore explains the relationship rather than relying on the broad Indonesian expression pengembalian uang alone. This scenario also gives the future needs analysis a concrete question: which repayment tasks, supporting documents, and English explanations occur in the intended organizations?

3.8. Technical Delivery and Accessibility Requirements

The proposed delivery model is a mobile-friendly reference resource that can retain a reviewed content package locally after installation or download. The paper prototype does not select a specific programming framework or claim offline behavior has been implemented. The intended requirement is that core definitions, examples, and practice items should remain available without continuous network access. Optional media can be downloaded separately so that users do not have to accept large files simply to read a short definition. The baseline display should use readable text, sufficient contrast, and headings that communicate structure. Information should not depend on color alone. A warning about invoice and receipt should appear as text, even if an interface also uses an icon or shading. Controls should have labels that remain meaningful when read by an assistive technology. Audio requires a text equivalent, while a fictional document image requires an accessible transcription of the information needed to answer the task.

Keyboard access matters even for a mobile-oriented resource because some users may use a laptop, an external keyboard, or an accessibility tool. Search, entry selection, related terms, and practice controls should be reachable in a predictable sequence. The visible focus indicator should remain clear. A future test should check whether the learner can return from a comparison screen without losing the original task. Accessibility review should cover the whole navigation path rather than a single attractive entry screen. A future technical test should deliberately interrupt connectivity, reload an entry, and inspect whether the latest approved content remains accessible. It should also test an unavailable audio file, a failed content update, and a search query with no result. The expected behavior must be specified before testing, such as keeping the definition readable when audio fails. These are proposed acceptance cases. No claim is made here that a software build has passed them.

3.9. Content Validation and Usability Evaluation Protocol

Content review should invite complementary expertise in finance, English pedagogy, and relevant local administration. The proposed review form asks whether each definition is accurate, its translation is appropriate, its example is understandable, and its caution addresses a meaningful confusion. Reviewers should be given the institutional scope and expected audience. They should also be able to mark an entry as requiring revision irrespective of its numerical rating if the wording could materially mislead a user about a transaction or document.

Yusoff describes systematic content validation and calculation of the content validity index [14]. A narrative review of content validity estimation likewise describes expert-panel ratings and indices such as the

content validity index and modified kappa [31]. A later study can adapt this method using a four-point relevance scale, with the two upper categories counted as agreement. The item-level index would be the number of reviewers endorsing relevance divided by the number rating that item. A scale-level average can summarize the item indices. The protocol should specify the panel size and decision criteria before ratings are collected. This article reports no computed index and does not claim expert validation.

Quantitative agreement should be accompanied by qualitative revision records. A strong average can conceal one inaccurate definition, and agreement about relevance does not establish linguistic clarity. Reviewer comments should identify the problem, proposed change, and reason for the decision. The revised entry should then be checked together with its linked practice item. If a definition changes from payment request to payment acknowledgment, the answer key and feedback must change as well. Content quality depends on consistent relationships across the record, not merely on a final list of headwords.

Usability testing should follow implementation of a functional prototype. Kumar, Chand, and Goundar's mapping study describes the value of user-based testing and common data collection methods in mobile learning evaluation [17]. A standardized questionnaire such as the System Usability Scale, described as the most widely used instrument for perceived usability [32], could be added to the observation protocol. For the proposed glossary, observation of realistic tasks should complement a questionnaire. Participants can be asked to find a term from an Indonesian query, compare two document functions, interpret a fictional amount due, and resume an entry after a simulated interruption. The observer records success, assistance, navigation errors, and the participant's explanation of what happened. Table 8 distinguishes proposed measures and their interpretation limits. Completion rates describe task performance under the test conditions, while satisfaction ratings describe perceptions.

Table 8. Proposed Measures and Interpretation Limits

Dimension	Proposed evidence	Supported interpretation	Limit
Content accuracy	Finance review with revision records	Definitions fit the stated scope	Does not establish learner understanding
Language clarity	Language review and comprehension probes	Wording is interpretable in tested tasks	May differ across proficiency levels
Technical reliability	Acceptance tests on specified devices	Functions meet tested requirements	Does not cover every device or future update
Usability	Task success, assistance, and observed errors	Users can perform tested actions	Does not establish vocabulary retention
Vocabulary recognition	Independent meaning selection items	Recognition at the testing time	Does not establish productive use
Recall and use	Recall items and scenario rubric	Retrieval and contextual performance	Requires reliable scoring and task coverage
Workplace transfer	Authorized observations and document review	Use in the observed work context	Does not alone establish causal governance change

Learning evaluation should use independent items rather than simply repeating the practice questions. Repeating an identical quiz may measure memory for an answer position or scenario wording. A test can use a new amount, a different fictional document layout, or another transaction while retaining the target distinction. Recognition, recall, and contextual use should each have an explicit scoring rule. If the study includes open responses, two trained raters can independently score a subset or the full set, depending on the approved design, and report agreement and disagreement resolution.

The contextual-use rubric should reward financial meaning as well as English clarity. For example, an explanation of a balance due receives credit for identifying the total charge, acknowledging the advance, and giving the remaining amount correctly. Minor grammatical errors should not outweigh correct financial interpretation unless they change the meaning. Conversely, a fluent sentence that treats an invoice as proof of payment should lose conceptual credit. This separation reflects the actual aim of the glossary: useful, accurate communication about financial tasks.

3.10. Learning Study Design and Analysis

A later learning study should specify its primary outcome before recruitment. One defensible outcome is improvement in contextual financial vocabulary use, measured through independent scenarios. Recognition and recall can serve as secondary outcomes. A delayed assessment is necessary if the study aims to make a retention claim. The timing should be justified in relation to the intended use and practical constraints. The proposed paper prototype does not choose a final retention interval because the length and frequency of future

training have not been established. Where feasible, a comparison condition should provide the same lexical content in a static format. This would help distinguish the contribution of interaction from the contribution of exposure to accurate content. Both groups should receive comparable training time and access to the same core terms. If one condition has more practice, more instructor attention, or different examples, the interpretation must acknowledge those differences. A comparison between unequal content packages cannot isolate interactivity as the cause of a learning difference.

Random assignment may be difficult when managers work in the same organization and share materials. If a quasi-experimental design is used, the study should report baseline differences and the procedures used to address them. If allocation occurs by village or training group, analysis must account for clustering rather than treating every response as independent. If there is only one group with pretest and posttest, the results can describe change but provide weaker support for attributing that change to the glossary. The design choice must match the intended causal claim. Sample planning should follow the chosen analysis, the smallest meaningful difference, expected score variability, attrition, and any group-level allocation. Formative convenience samples require caution when generalizing beyond the participants.

Analysis should report descriptive results and uncertainty, not only a significance decision. For each outcome, the study should identify the score range, distribution, number assessed, and amount of missing data. Effect estimates and confidence intervals can help readers judge practical magnitude when the design supports them. For small or skewed samples, an analysis suited to the data may be more defensible than applying a familiar test mechanically. Statistical assumptions and deviations from the planned analysis should be explained. Qualitative analysis can investigate why an entry or activity succeeds or fails. Interviews and observed explanations may reveal that a translation is unfamiliar, an example does not match local work, or a task category uses the wrong organizational label. Coding should distinguish language difficulty, conceptual confusion, navigation difficulty, and institutional mismatch. A revision based on such evidence can be justified more precisely than a generic statement that participants wanted an improved application. Reporting should include the analysis procedure without inventing saturation or interrater reliability.

3.11. Failure Analysis and Content Governance

Table 9 identifies anticipated failure cases and proposed responses. These cases guide future testing; their frequency in the intended population is unknown.

Table 9. Anticipated Failure Cases and Proposed Responses

Failure Case	Why It Matters	Proposed Response
Invoice treated as receipt	User may incorrectly confirm payment.	Comparison activity and explicit document status note
Allocation treated as spending	Plan is misrepresented as an actual transaction.	Contrast budget allocation with expenditure
Cash balance treated as surplus	Different financial measures are merged.	Require entity, period, and reporting basis
Translation treated as a legal rule	Local procedure may be misapplied.	Context label and reviewer approval before release
Search returns a broad synonym first	User may select the wrong concept.	Rank exact matches and label suggestions
Offline version becomes outdated	Users may retain a corrected definition.	Visible version and controlled update notice
Practice success called competence	Evidence is overstated.	Report recognition, recall, and use separately

Content governance begins with responsibility for changes. A later project should identify who can propose an entry, who reviews its finance meaning, who reviews its English, and who approves publication. The same person may fill more than one role in a small project, but the responsibilities should remain explicit. A suggested term should not become public merely because it has been submitted. New content requires the same checks as the original starter set, including its associated examples and learning items.

Version records should preserve the reason for a change. A typographical correction differs from a revised definition, and a revised definition can change the correct answer to a task. Keeping a change log allows instructors to determine which materials need replacement. It also supports reproducibility in a later learning study: readers need to know which content version participants actually used. An application version alone is insufficient if the entry package changes independently of the software.

3.12 Comparison With Prior Studies

This subsection compares the design outputs with prior studies in four respects: lexical content, entry structure, evidence base, and evaluation. Table 10 summarizes the comparison. It concerns design alignment, not effectiveness, because this study has no learner data.

Table 10. Comparison of the Design Outputs With Prior Studies

Aspect	Prior Studies	Relation to the Present Design	Status of the Interpretation
Terminology difficulty	Accounting students find financial terms in reports, auditing, and budgeting difficult [21]	Aligned: the starter set emphasizes budget, budget realization, variance, audit, and financial report	Relevance for village managers is Pending
Bilingual entry structure	English word, Indonesian translation or description, bilingual example sentences, and pictures [15]	Aligned on core fields; extended with institutional context, caution, learning prompt, and review fields	Whether the extra fields help or burden users is Pending
Visual support	Students asked for thematic illustrations in a picture dictionary [19]; gloss effects vary by proficiency and test format [10], [29]	Divergent: definition and example come first, and images and audio are optional	A testable hypothesis, Pending
Needs analysis	Surveys, focus groups, and interviews with students, lecturers, and practitioners [16], [19]; community English studies [5]-[7]	Divergent: scenarios derive from literature and the authors' analysis, with no needs-analysis data	Local needs analysis is Pending
Spaced and mobile practice	Spacing, mobile vocabulary, and digital reading syntheses [20], [25], [26], [28]	Aligned as design inference: optional review and links from terms to short texts	Effects are not transferred; Pending
Validation and testing	Expert validation and field testing [15]; content validity [14]; usability mapping [17]	Aligned in sequence; this study stops before validation	Pending
Governance	Transparency, human resource quality, internal control, and access to information matter; technology was not significant in two models [2], [33]; limited public access [3]	Consistent: the glossary targets communication, not oversight	No governance effect is claimed

The emphasis of the starter set on reporting and budgeting terms is consistent with evidence that accounting students in Indonesia struggle with English financial terminology [21]. The agreement is limited: those learners are university students, and [21] evaluated a project-based learning model, not a glossary. It supports the plausibility of the lexical focus, not the effectiveness of this design. The entry structure resembles the English–Indonesian e-glossary of [15], which was developed through preliminary study, product development, expert validation, and field testing. The present schema adds institutional context, a usage caution, and links to a learning prompt and review fields. These additions respond to a feature of the target domain: terms such as revenue, asset, and own source revenue change meaning with the reporting entity, which a general ESP glossary need not handle.

The designs diverge on visual support. Students in [19] wanted thematic illustrations, whereas the present design gives priority to definition and example, in line with syntheses showing that gloss effects vary [10], [29], including a meta-regression that found no significant difference among textual, pictorial, and auditory gloss modes [22], and because many financial concepts are abstract. The divergence may reflect domain (concrete aviation procedures versus abstract finance) or population. It is a hypothesis to test, not a finding. Whether adding pictures to glosses enhances vocabulary uptake has been examined experimentally for general vocabulary [34]; a comparable test with financial terms would be needed. The most consequential difference concerns evidence. The needs analyses in [16] and [19] collected field data, and community English studies in tourism settings [5]-[7] likewise rest on field data. The present study derives its scenarios from the literature, so it cannot say which terms local managers need. It can specify the evidence needed to decide, which is why the work is framed as conceptual development.

Practice and evaluation proposals draw on syntheses of spacing, mobile vocabulary learning, and digital reading [20], [25], [26], [28]. These report positive but condition-dependent outcomes, so their effect sizes are

not transferred and the review intervals remain Proposals. The sequence of expert validation followed by user testing follows; unlike [15], this study reports no effectiveness evidence. The large improvement reported for an ESP model in accounting classes [21] concerns a different intervention and population and is not an expected effect of the glossary. Governance studies indicate that transparency, human resource quality, and internal control matter for village accountability, and that information technology was not significant in two models [2], [33]. Limited public access to information about village-owned enterprises [3] shows that clear explanations are one part of disclosure only. The findings are therefore consistent with treating the glossary as support for communication, not as a mechanism of accountability.

3.13 Novelty of the Study

Novelty is assessed by comparison with the closest studies in Table 11. Individual features such as bilingual entries, example sentences, mobile delivery, needs analysis, and spaced practice are not new; they appear in [15], [19], [21] and the literature in Section 1. The contribution lies in how they are configured for a domain and audience not covered in the reviewed studies, and in how claims are managed. Three elements together distinguish the study. First, an institution-aware entry schema that records context and cautions for terms whose meaning depends on the reporting entity. Second, a task and document route linked to financial communication scenarios, with feedback that names the financial distinction at issue. Third, a status register and pre-specified evidence plan that separate outputs, proposals, and pending results. Whether these elements improve accuracy, usability, or learning is Pending, so the claim is one of specification, not of demonstrated superiority.

Table 11. Comparison With the Closest Prior Studies

Dimension	Rohani and Suyono [15]	Suhaimi et al. [21]	Ginusti and Ariebowo [19]	This study
Target users	Students in commerce and engineering departments	Accounting students (120, three universities)	Air transport management students (141)	Tourism village financial managers (intended; no participants)
Product	Android bilingual e-glossary	ESP-based learning model with modules and project work	Picture dictionary (needs-analysis stage)	Glossary specification (paper prototype)
Entry or content focus	English word, Indonesian translation or description, bilingual examples, pictures	Reports, financial statements, budgeting, audit scenarios	Vocabulary lists, pronunciation, definitions, word class, illustrations (requested by students)	English term, Indonesian equivalent, plain definition, example, institutional context, caution, learning prompt, review fields
Evidence reported	Expert validation and field testing	Significant improvement ($p < .001$, $d = 1.21$)	Needs-analysis findings	Design outputs only
Distinctive element here	Not applicable	Not applicable	Not applicable	Institution-aware schema; task route with distinction-focused feedback; status register and evidence plan

3.14. Educational Contribution and Regional Application

The central educational contribution is the integration of financial concept explanation with role-sensitive English use. A general bilingual word list can support translation, but it may leave the learner unable to explain a document's function or the relationship between amounts. The proposed glossary makes these relationships visible through definitions, examples, comparisons, and short scenarios. Its distinctive element is this coordination, rather than the claim that any individual feature is technically new. This approach can support collaboration between economics education and English education. Economics educators can examine conceptual boundaries and document interpretation. English educators can examine vocabulary selection, grammar, intelligibility, and scaffolded production. Joint review helps preserve financial accuracy while making the explanations accessible.

For North Sulawesi, the application should begin with a small number of partner organizations whose tasks and permissions are clearly defined. Published studies establish useful regional context, but they do not

supply a complete language-needs profile. A later team should first identify the specific organization, the financial manager's responsibilities, and the communication occasions involving English. If English is used rarely, a concise reference tool may be more appropriate than an intensive course. If users frequently explain payments or reports to external partners, structured practice may have greater relevance.

The glossary should also respect Indonesian as the principal language of many internal administrative activities. English support can help with particular communication needs without implying that effective financial governance requires replacing local working languages. A bilingual resource lets users connect an English expression to a familiar conceptual explanation. Local terminology and cultural practices can be represented in examples after review, but their inclusion should follow the users' actual tasks rather than decorative references to regional identity.

The governance literature also limits the expected impact. Damayanti, Hanafi, and Sutapa report relationships between village apparatus competence, participation, internal control, sanctions, and accountability, while the use of information technology was not significant in their model [33]. Their study does not evaluate a glossary, but it reinforces a cautious interpretation of digital interventions. The present design may support accurate communication; it cannot alone establish improved oversight, reduced misuse of funds, or greater public trust.

Institutional outcomes depend on financial procedures, incentives, access to information, and organizational responsibility as well as individual understanding. If a manager learns the English word accountability but the organization does not publish understandable reports, the vocabulary gain does not resolve the access problem. If a manager correctly explains receipt but supporting evidence is poorly maintained, a language resource does not repair the records. The proposed glossary should therefore sit alongside appropriate finance training and governance arrangements rather than carry claims beyond its educational function.

Because the findings are design findings, generalization here is analytic and conditional, not statistical: the question is which design principles may be carried to similar settings, not which outcomes will be replicated. Five principles are proposed. (1) Where a term's meaning depends on the institution, an entry should state its scope and a caution; the schema in Section 3.2 and the failure cases in Section 3.11 motivate this principle. (2) Task and document routes should complement alphabetical access, consistent with the emphasis on role-based provision in tourism village English studies [5], [6]. (3) Feedback should name the distinction at issue, not only mark an answer correct or incorrect. (4) Recognition, recall, and contextual use should be assessed separately, consistent with the variation reported for glossing [10], [29]. (5) Claims should be tied to a status register so that outputs, proposals, and pending results are not conflated.

These principles may be transferred to other tourism villages, village-owned enterprises, and community organizations in Indonesia that communicate financial information in English, and possibly to other institution-dependent domains such as administrative or legal English. Transfer is conditional on a local needs analysis, on the reporting framework of the organization, and on English being used often enough to justify a dedicated resource. The principles are Proposals. No participants, sites, or outcomes support their empirical generalization, and they should be treated as hypotheses for the next stage.

3.15. Limitations and Next Development Stage

The study has four main limitations. First, it is conceptual and provides no primary needs-analysis data from the intended users. The relevance of the starter terms remains provisional. Second, the targeted review does not establish comprehensive coverage of the literature or comparative superiority over existing products. Third, the paper prototype specifies interaction but does not demonstrate software reliability or accessibility. Fourth, the definitions and translations have not undergone independent expert validation. These limits restrict the claims that can be drawn from the design outputs.

The thirty-term scope is illustrative. It omits many concepts that may become relevant, such as receivables, depreciation, restricted funds, or more specialized government reporting expressions. Their exclusion is not a judgment that users do not need them. A later needs analysis should prioritize terms by actual task relevance, frequency in authorized documents (corpus use for second language vocabulary learning has been synthesized meta-analytically [35]), consequences of confusion, and learner difficulty. These criteria should be reported, allowing readers to see why the revised set differs from the present starter package.

The next development stage should gather local evidence, revise the entry set, obtain content review, and implement a small functional version. Only then should observed usability and learning data be collected. The sequence can be iterative: a failed comprehension probe may require a definition change before the next interface test. The project should document these revisions so that a later report describes what participants actually used and what changed after their feedback.

The study has five main limitations. First, it is conceptual and provides no primary needs-analysis data from the intended users, so the relevance of the starter terms is Pending. Second, the targeted review does not establish comprehensive coverage of the literature or superiority over existing products, and the comparisons in Section 3.12 rest on abstracts and published summaries, not on reanalysis of those studies. Third, the paper

prototype specifies interaction but does not demonstrate software reliability or accessibility. Fourth, the definitions and translations have not undergone independent expert validation, and the scenario review was conducted by the authors only. Fifth, the generalization in Section 3.14 is analytic and conditional. These limits restrict the claims to the Output status in Table 3.

4. CONCLUSION

In answer to RQ1, the article derived an evidence-informed requirement structure and a task-based framework from the literature, and drafted thirty bilingual entries across six working domains (Sections 3.1 to 3.3). The entries are unreviewed drafts, and their local relevance is yet to be determined. In answer to RQ2, the article specified one coherent resource: a record schema with institutional context and usage cautions, search and task routes, separate activities for recognition, recall, and contextual use, explanatory feedback, and delivery requirements, illustrated by three model entries and three worked scenarios (Sections 3.4 to 3.8). These are specifications, not implemented or tested features. In answer to RQ3, the article defined the evidence needed to assess content accuracy, usability, vocabulary learning, and workplace relevance through content validation, observed user tasks, and a learning study design (Sections 3.9 to 3.11). None of this protocol has been carried out. Compared with earlier bilingual glossaries and ESP studies, the distinctive element of this work is its configuration for a new audience and domain: institution-aware entries, task-based access, and feedback that names the financial distinction at issue, together with an explicit separation of outputs, proposals, and pending results. The study does not establish user demand, usability, vocabulary gains, or improved financial governance, and no such benefit should be inferred from the design alone.

Recommendations for use are limited. The draft entries should serve only as supplementary educational material, not as an accounting dictionary, legal reference, or official translation, and only after a formal term-review process involving finance, English, and local village administration reviewers. Village governments, tourism offices, and training providers that consider such a resource could support three preparatory actions: commissioning a needs analysis in a small number of partner organizations, establishing a review and version-control procedure for English equivalents of village finance terms, and piloting any resulting version with clearly defined users before wider adoption. Adoption decisions should wait for the evidence described below. Future research should proceed in this order. First, a local needs analysis should examine roles, authorized documents, observed communication, and the terms users find difficult. Second, independent experts should validate the definitions and equivalents, for example through a content validity index. Third, a functional prototype should be tested with prospective users for usability, accessibility, and offline behavior. Only then should a learning study use independent items and a suitable comparison condition. Claims about retention and workplace transfer should follow separate assessments, not satisfaction or practice completion.

ACKNOWLEDGEMENTS

The author expresses gratitude to all those involved in this research, whose contributions ensured that the study proceeded smoothly.

REFERENCES

[1] V. A. Kantale, A. T. Poputra, and H. R. N. Wokas, “Analisis akuntabilitas pengelolaan dana desa dalam meningkatkan pemberdayaan masyarakat di Desa Darunu Kecamatan Wori Kabupaten Minahasa Utara: Studi kasus pelaksanaan pembangunan wisata Mangrove,” *Riset Akuntansi dan Portofolio Investasi, vol. 3*, no. 2, pp. 558–569, 2025, doi: 10.58784/rapi.392.

[2] A. Salindeho, A. Gazali, and P. J. Pesak, “Pertumbuhan desa wisata: Digitalisasi BUMDes, transparansi laporan keuangan, kualitas sumber daya manusia,” *Solusi, vol. 24*, no. 2, pp. 192–212, 2026, doi: 10.26623/slsi.v24i2.14065.

[3] R. J. Kalangie, S. Sampe, and N. Kumayas, “Transparansi pengelolaan Badan Usaha Milik Desa di Kecamatan Langowan Barat Kabupaten Minahasa,” *Governance, vol. 3*, no. 2, 2023.

[4] P. Permatasari et al., “Village fund management and reporting systems: Are they accountable?,” *Transforming Government: People, Process and Policy, vol. 18*, no. 4, pp. 512–528, 2024, doi: 10.1108/TG-07-2023-0098.

[5] C. Manara and Yanti, “Exploring the English language needs of a community-based tourist village in an Indonesian context,” *Scope: Journal of English Language Teaching, vol. 9*, no. 1, pp. 85–93, 2024, doi: 10.30998/scope.v9i1.23047.

[6] H. A. Wibowo, M. Nurharjanti, and R. J. Vijaya, “A need analysis for local tour guides at Jagalan Tourism Village in Yogyakarta,” *Journal of Innovation Research and Knowledge, vol. 4*, no. 7, pp. 4575–4580, 2024, doi: 10.53625/jirk.v4i7.9093.

[7] R. Suprina, N. Rahayu, and V. A. Sembiring, “Exploring English communication needs in a tourism village: What do the local communities say?,” *Journal of Business on Hospitality and Tourism, vol. 9*, no. 1, pp. 74–80, 2023, doi: 10.22334/jbhost.v9i1.432.

[8] A. Yu and G. Trainin, "A meta-analysis examining technology-assisted L2 vocabulary learning," *ReCALL, vol. 34*, no. 2, pp. 235–252, 2022, doi: 10.1017/S0958344021000239.

[9] K. Okumuş Dağdeler, "A systematic review of Mobile-Assisted Vocabulary Learning research," *Smart Learning Environments, vol. 10*, Art. no. 19, 2023, doi: 10.1186/s40561-023-00235-z.

[10] H. S. Mahdi, M. A. Mohsen, and M. Almanea, "Multimedia glosses and second language vocabulary learning: A second-round meta-analysis," *Acta Psychologica, vol. 248*, Art. no. 104341, 2024, doi: 10.1016/j.actpsy.2024.104341.

[11] J. Sweller, J. J. G. van Merriënboer, and F. Paas, "Cognitive architecture and instructional design: 20 years later," *Educational Psychology Review, vol. 31*, no. 2, pp. 261–292, 2019, doi: 10.1007/s10648-019-09465-5.

[12] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *Journal of Management Information Systems, vol. 24*, no. 3, pp. 45–77, 2007, doi: 10.2753/MIS0742-1222240302.

[13] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Quarterly, vol. 28*, no. 1, pp. 75–105, 2004, doi: 10.2307/25148625.

[14] M. S. B. Yusoff, "ABC of content validation and content validity index calculation," *Education in Medicine Journal, vol. 11*, no. 2, pp. 49–54, 2019, doi: 10.21315/eimj2019.11.2.6.

[15] S. Rohani and A. Suyono, "Developing an Android-based bilingual e-glossary application of English for Specific Purposes (ESP)," *English Language Teaching Educational Journal, vol. 4*, no. 3, pp. 225–234, 2021, doi: 10.12928/eltej.v4i3.5209.

[16] A. Nugroho, "English material needs of accounting students: An English for Specific Purposes approach," *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature, vol. 8*, no. 1, pp. 1–15, 2020, doi: 10.24256/ideas.v8i1.1239.

[17] B. A. Kumar, S. S. Chand, and M. S. Goundar, "Usability testing of mobile learning applications: A systematic mapping study," *International Journal of Information and Learning Technology, vol. 41*, no. 2, pp. 113–129, 2024, doi: 10.1108/IJILT-03-2023-0029.

[18] B. Pang, F. Abd Majid, and G. Narayanan, "Systematic literature review on needs analysis in ESP within higher education: Implications for future research," *Malaysian Journal of Social Sciences and Humanities, vol. 10*, no. 1, 2025, doi: 10.47405/mjssh.v10i1.3162.

[19] G. N. Ginusti and T. Ariebowo, "Developing an Air Travel Picture Dictionary in enhancing vocabulary mastery among air transport management students: A need analysis," *Voices of English Language Education Society, vol. 7*, no. 3, pp. 572–586, 2023, doi: 10.29408/veles.v7i3.24039.

[20] S. K. Kim and S. Webb, "The effects of spaced practice on second language learning: A meta-analysis," *Language Learning, vol. 72*, no. 1, pp. 269–319, 2022, doi: 10.1111/lang.12479.

[21] I. Suhaimi, C. Hana, and N. A. Sudibyo, "ESP-based English learning model for enhancing financial competence among accounting students: A design-based research study," *Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran, dan Pembelajaran, vol. 12*, no. 1, pp. 40–50, 2026, doi: 10.33394/jk.v12i1.19346.

[22] A. Yanagisawa, S. Webb, and T. Uchihara, "How do different forms of glossing contribute to L2 vocabulary learning from reading? A meta-regression analysis," *Studies in Second Language Acquisition, vol. 42*, no. 2, pp. 411–438, 2020, doi: 10.1017/S0272263119000688.

[23] O. O. Adesope, D. A. Trevisan, and N. Sundararajan, "Rethinking the use of tests: A meta-analysis of practice testing," *Review of Educational Research, vol. 87*, no. 3, pp. 659–701, 2017, doi: 10.3102/0034654316689306.

[24] T. Uchihara, S. Webb, and A. Yanagisawa, "The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies," *Language Learning, vol. 69*, pp. 559–599, 2019, doi: 10.1111/lang.12343.

[25] Y. Zhou and M. Zhou, "A meta-analysis on mobile-assisted vocabulary learning: Do mobile applications help?," *ReCALL, vol. 38*, no. 1, pp. 75–93, 2026, doi: 10.1017/S0958344025100335.

[26] C. Dan, L. Ismail, A. B. Razali, and L. Dandan, "A meta-analysis of the existing studies on effects of mobile learning on vocabulary acquisition," *International Journal of Instruction, vol. 18*, no. 3, pp. 765–784, 2025.

[27] B. Wisniewski, K. Zierer, and J. Hattie, "The power of feedback revisited: A meta-analysis of educational feedback research," *Frontiers in Psychology, vol. 10*, Art. no. 3087, 2020, doi: 10.3389/fpsyg.2019.03087.

[28] T. Zhu, Y. Zhang, and D. Irwin, "Second and foreign language vocabulary learning through digital reading: A meta-analysis," *Education and Information Technologies, vol. 29*, no. 4, pp. 4531–4563, 2024, doi: 10.1007/s10639-023-11969-1.

[29] M. A. Mohsen, H. S. Mahdi, and R. Alkhammash, "Multimedia glosses and their impact on second language vocabulary acquisition: Insights from a meta-analysis and document co-citation analysis," *Innovation in Language Learning and Teaching, vol. 18*, no. 2, pp. 109–124, 2024, doi: 10.1080/17501229.2023.2236084.

[30] S. Webb, A. Yanagisawa, and T. Uchihara, "How effective are intentional vocabulary-learning activities? A meta-analysis," *The Modern Language Journal, vol. 104*, no. 4, pp. 715–738, 2020, doi: 10.1111/modl.12671.

[31] E. Almanasreh, R. Moles, and T. F. Chen, "Evaluation of methods used for estimating content validity," *Research in Social and Administrative Pharmacy, vol. 15*, no. 2, pp. 214–221, 2019, doi: 10.1016/j.sapharm.2018.03.066.

[32] J. R. Lewis, "The System Usability Scale: Past, present, and future," *International Journal of Human–Computer Interaction, vol. 34*, no. 7, pp. 577–590, 2018, doi: 10.1080/10447318.2018.1455307.

[33] N. V. Damayanti, R. Hanafi, and Sutapa, "The accountability of village fund management: Study on villages in the sub-district of Dempet," *Journal of Advanced Multidisciplinary Research, vol. 5*, no. 1, pp. 60–71, 2024, doi: 10.30659/jamr.5.1.60-71.

[34] F. Boers, P. Warren, L. He, and J. Deconinck, "Does adding pictures to glosses enhance vocabulary uptake from reading?," *System, vol. 66*, pp. 113–129, 2017, doi: 10.1016/j.system.2017.03.017.

[35] H. Lee, M. Warschauer, and J. H. Lee, "The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis," *Applied Linguistics, vol. 40*, no. 5, pp. 721–753, 2019, doi: 10.1093/applin/amy012.