INTERACTION: Jurnal Pendidikan Bahasa Vol. 13, No. 3; September 2026
P-ISSN: 2406-9558; E-ISSN: 2406-9566

Feasibility and Preliminary Vocabulary Learning Outcomes of a Thematic Digital Visual Dictionary for Mandailing, Indonesian Based on Local Wisdom

Eko Firman Susilo
ekofirmansusilo949@gmail.com

Anna Leli Harahap
annaleliharahap25@gmail.com

Asri Yulianda
asriyulianda23@gmail.com

Shopiah Anggraini Rambe
anggrainishopiah@gmail.com

Nurfadhillah Siregar
fadhillahnur137@gmail.com

Universitas Al Washliyah

ABSTRACT

Indonesia's reading literacy performance, as reported in the OECD's PISA 2022 results, remains below the OECD average, and limited vocabulary mastery among students in bilingual communities such as the Batak Mandailing is often cited as a contributing factor. At the same time, everyday use of the Mandailing language among younger community members appears to be declining. This study reports the development, expert validation, and a preliminary single group trial of a Thematic Digital Visual Dictionary for Mandailing, Indonesian based on local wisdom, developed as a vocabulary learning medium for elementary school students using the Research and Development (R&D) approach with the ADDIE model (Analysis, Design, Development, Implementation, Evaluation). The product was reviewed by three experts, one material expert, one media expert, and one language expert, who rated it Highly Feasible, with a combined mean validation score of 3.51 on a 4 point scale. In a one group, one session pretest, posttest trial conducted 1 to 5 July 2026, all 25 students of the Grade 5 class (15 female, 10 male; age 10 to 11 years) at SD UPTD 04 Simatahari (population 156) completed a Mandailing thematic vocabulary test immediately before and after an approximately 60 minute period of independent product use, then completed a user response questionnaire. Users rated the product positively for practicality and usefulness (mean = 3.48, Highly Feasible category). Mean thematic vocabulary scores rose from 54.12 (pretest) to 79.40 (posttest), with an average Normalized Gain (N Gain) of 0.56 (moderate category; Hake, 1998). Because the design used a single group, one immediate posttest, and no comparison condition, these results are best read as preliminary evidence of product feasibility and a short term thematic vocabulary gain, not as

evidence of broader language literacy improvement or language revitalization; controlled and longer term studies are needed before such claims can be supported.
Keywords: Digital Visual Dictionary; Batak Mandailing; Local Wisdom; Thematic Vocabulary Learning; Society 5.0

INTRODUCTION

Advances in digital technology increasingly call for education to offer more interactive, engaging, and contextual learning media, including for language learning at the elementary school level. In Indonesia, elementary students' reading literacy remains a national concern: the OECD's PISA 2022 results place Indonesia's mean reading score well below the OECD average (OECD, 2023). Limited vocabulary mastery is frequently identified as one contributor to weak literacy outcomes, particularly for students in bilingual or multilingual regions such as the Batak Mandailing community. At the same time, use of Mandailing and other local languages is reported to be declining; the Badan Pengembangan dan Pembinaan Bahasa [Language Development and Fostering Agency] (2024) lists more than 25 local languages as endangered absent revitalization efforts. The link between these national- and system level indicators and the specific situation at the target school is inferential rather than directly demonstrated in this study.

One potentially useful medium is the thematic digital visual dictionary. Prior studies of dictionary based learning media have focused mainly on general bilingual dictionaries, and development work specifically grounded in Batak Mandailing local wisdom for direct classroom use remains limited; most existing products have also not combined visual content, local cultural context, and digital delivery within a single, integrated tool. This study responds to that gap by developing a thematic digital visual dictionary for Batak Mandailing, Indonesian intended to function both as a vocabulary learning tool and, longer term, as a resource that supports local language and cultural preservation.

This study is motivated by three considerations: (1) elementary students' reading literacy and vocabulary mastery, particularly among bilingual Batak Mandailing students, motivate interest in more interactive, digitally delivered learning media; (2) the reported decline in everyday Mandailing use motivates interest in learning resources that can be used directly in elementary classrooms; and (3) existing dictionary resources have not, to our knowledge, integrated visual elements, local cultural context, and digital delivery in one product. These considerations motivate, but, given the present design, do not by themselves establish, the value of the product developed here; the actual contribution this study can support is described precisely in the research questions below.

This study addresses three research questions, aligned with the outputs of the ADDIE model: 1) What thematic vocabulary needs do students, teachers, and community informants report for Mandailing, Indonesian learning, and how did

these needs inform the product's content and design? 2) Do material, media, and language experts judge the developed product's content, media, and language to be feasible for elementary school use? 3) Do users perceive the product as practical and useful, and is there a measurable short term change in Mandailing thematic vocabulary scores after one guided use session?

The primary outcome of this study is Mandailing thematic vocabulary score, measured immediately before and after a single approximately 60 minute session of product use, in one group with no comparison condition. Expert validation and user response scores are secondary, process level outcomes describing product feasibility and perceived practicality; they are not treated as evidence of learning impact. This study does not measure, and therefore does not claim, improvement in general reading or language literacy, strengthened cultural identity, or language revitalization; these terms appear later only as motivating background and as directions for future research, not as findings of the present trial.

Terminology is used consistently as follows in this manuscript: “Mandailing” and “Batak Mandailing” refer interchangeably to the same ethnolinguistic group and language, following the usage in the cited sources; “local language” refers to Mandailing specifically, not to local languages generally; and the 25 Grade 5 pupils who took part in the trial are referred to as “students” or “participants” throughout, not as “respondents,” “speakers,” or “learners,” terms which imply a broader or self selected population that this study did not sample.

Society 5.0 (Fukuyama, 2018) is invoked here not simply as a label for “using digital technology,” but because it foregrounds using technology deliberately to address a specific social problem, in this case, the reported decline in intergenerational transmission of Mandailing vocabulary, rather than technology adoption for its own sake. Under this framing, the digital format is a means of extending a culturally grounded learning resource to students who already use smartphones daily, not the study's main contribution in itself; the ADDIE based development process and its preliminary evaluation, described below, are the empirical basis for the study's claims. The contribution of this study is not any single feature in isolation, digital delivery, visual illustration, native speaker audio, thematic organization, and local wisdom curation all appear individually in prior work, but their combination for Mandailing, Indonesian specifically, paired with a preliminary pretest, posttest trial. Table 1 compares the present product against the closest prior products feature by feature, to make this basis for the novelty claim explicit rather than asserted.

LITERATURE REVIEW

Mandailing-Language Status and Educational Need

Aprilia et al. (2025) examined Mandailing-language maintenance amid globalization and reported that everyday use among younger community members

continues to decline due to the dominance of Indonesian and foreign languages; their study was descriptive/qualitative and did not test an intervention, leaving open what kind of resource might support maintenance in a classroom setting. The Badan Pengembangan dan Pembinaan Bahasa (2024) reports more than 25 Indonesian local languages as endangered absent revitalization effort, a system level statistic that motivates but does not by itself demonstrate the specific need at SD UPTD 04 Simatahari (*see Introduction*). Together, these sources establish that Mandailing maintenance is a recognized concern, while indicating a gap for classroom based, evaluable interventions rather than documentation or description alone, the gap this study addresses.

Visual and Thematic Dictionary Design

Tarp and Gouws (2023) argue that digital lexicography differs functionally from print dictionary making, and that digital tools should be evaluated on the specific access, retrieval, and multimedia affordances they add rather than treated as a straightforward digitization of print convention; this framing supports evaluating the present product on its thematic organization, audio, and low data web access specifically, rather than on "digital ness" in general. Rahmawati and Suliyatun (2021) developed a printed Arabic, Indonesian pocket dictionary using a Research and Development approach; the product was judged feasible by material and media experts and received a positive response from students in small scale and field trials, but the study used no digital or audio visual elements and no learning outcome measure, limiting its comparability on both the product and the evaluation side. Haryani and Yustha (2021) developed a bilingual illustrated children's dictionary for Dayak Ma'anyan, English using a similar R&D model with expert validation in the Highly Feasible category; the product added illustration but, like the previous study, remained a conventional printed medium without digital search, audio, or a measured learning outcome.

Local Wisdom Based Language Learning

Bogar et al. (2025) developed an Android based electronic dictionary for the Napan language using an experimental/prototyping approach with native speaker involvement in data collection; the application supported search and cross community communication but did not integrate thematic cultural organization and did not include a pretest, posttest or other measured learning outcome. Tambaip et al. (2020) developed an Android based Muyu, Indonesian, English dictionary using a waterfall development method with a sample of 40 university level Muyu students; the study measured user preference and attitude toward the application rather than a vocabulary or literacy outcome, and targeted young adult rather than elementary school users. Across these studies, local wisdom or Indigenous language content is consistently paired with either a printed format or an unmeasured digital format; none combines thematic local wisdom curation, native speaker audio, and a measured elementary school vocabulary outcome, which is the specific combination this study reports on.

Evaluating Digital Learning Media

Sianturi et al. (2019) applied a child-centered design process, involving children directly in iterative usability testing, to a Batak/Indonesian/English language introduction application for children, and reported that iterative testing with the target age group shaped both content and interface decisions; this is a methodological model for usability evaluation that the present study's single session, end of use questionnaire does not fully replicate (see Usability and Practicality below). Fauzi and Rosliyah (2020) reported that a web based thematic dictionary was perceived positively by learners of beginner Japanese vocabulary, easing word search and supporting independent learning motivation, but relied on perception data only, without a learning outcome measure, a limitation shared, in part, by the user response component of the present study. Taken together, this literature indicates that (a) local language dictionary products are usually evaluated on feasibility/perception rather than measured learning outcomes, and (b) where child-centered usability methods exist, they involve iterative testing beyond what a single session evaluation can offer; the present study's pretest, posttest component addresses gap (a) preliminarily, while gap (b) is acknowledged as a limitation (see Conclusion and Limitations).

Conceptual Framework

Rather than treating “language literacy” as a single outcome, this study distinguishes proximal outcomes that the pretest, posttest can plausibly detect from longer term cultural aims that motivate the product but were not measured. Table 2 sets out this logic feature by feature: each product feature is linked to the learner activity it is intended to prompt, the proximal vocabulary outcome that activity is expected to produce, and a longer term cultural aim that the feature may eventually support but that this study does not test.

Table 1. Conceptual Framework

Product Feature	**Learner Activity**	**Proximal Vocabulary Outcome**	**Longer Term Cultural Aim**
Thematic categorization + illustration	Browsing/searching a theme; associating image with word	Faster recognition and recall of thematic vocabulary	Familiarity with culturally organized vocabulary domains
Native speaker pronunciation audio	Listening and repeating pronunciation	Correct pronunciation and sound, meaning association	Exposure to an authentic Mandailing dialect model
Local wisdom content (customs, kinship, food, house, clothing, symbols)	Encountering vocabulary embedded in cultural context	Contextualized rather than isolated word learning	Potential, unmeasured contribution to cultural identity and language maintenance
Self assessment quiz	Answering quiz items after browsing	Immediate feedback on recall	Sustained, self directed re engagement with the product (not measured here)

Local wisdom, in this framework, refers to a community's values, knowledge, and cultural character, formed through sustained interaction with the environment and transmitted across generations (Ayatrohaedi, 1986); it is operationalized here as the six thematic content categories described in Development Stage Results, not as a measured psychological construct. Society 5.0 (Fukuyama, 2018) functions in this framework as a rationale for the middle right column of Table 2, using digital technology deliberately to support a specific social aim (local language maintenance), rather than as a claim that the product itself achieves a Society 5.0 outcome; only the proximal, left hand columns of Table 2 are addressed by the evidence reported in Results.

As a development framework, this study uses the ADDIE model, Analysis, Design, Development, Implementation, Evaluation, a systematic, cyclical instructional design model in which each stage produces a defined output before the next stage begins (Branch, 2009). The model structures how the features in Table 2 were identified (Analysis/Design), built and validated (Development), and evaluated (Implementation/Evaluation); the traceability matrix in the Method section below documents this process stage by stage.

METHOD

Design and Sample

This study used the Research and Development (R&D) approach with the ADDIE model, Analysis, Design, Development, Implementation, and Evaluation (Sugiyono, 2021; Branch, 2009), to design, develop, and evaluate a technology based learning medium. The evaluation component that was actually conducted was a one group, one shot pretest, posttest pre experimental design (Sugiyono, 2021), carried out with a single class at Sekolah Dasar UPTD 04 Simatahari and no control or comparison class. The five ADDIE stages were carried out as follows: Analysis, 20 April, 7 May 2026; Design, 11 May, 8 June 2026; Development, 22 June, 30 June 2026; Implementation, 1 to 5 July 2026; and Evaluation, 10 to 27 July 2026. One round of revision occurred, during the Development stage, following expert validation; the revision consisted of corrections to the product's visual display and to spelling/orthography, after which the revised version (draft 2) was finalized as the product used in the trial. Because this is a pre experimental, one group design without random assignment or a comparison condition, pretest-to-posttest changes reported in the Results should be read as a preliminary, within group association rather than a controlled estimate of the product's causal effect; this limitation is discussed further below.

The population comprised 156 students at Sekolah Dasar UPTD 04 Simatahari, across all grade levels. The Grade 5 class (n = 25; 15 female, 10 male; age range 10 to 11 years) was selected as the trial class through purposive selection of Grade 5 as the target grade, combined with total (census) sampling of that entire class, the school's sole Grade 5 class, rather than a random draw from a larger pool; no random sampling step was used at any stage. No inclusion or exclusion criterion based on Mandailing-language background was applied; all students in the Grade 5 class were included regardless of home language background. Attendance across the trial was reported as high, with no dropouts among the 25 analyzed participants. All 25 Grade 5 students participated in the needs related classroom observation, the pretest, posttest trial, and the user response questionnaire, conducted during the Implementation window of 1 to 5 July 2026. School authorization for the study was obtained from SD UPTD 04 Simatahari (see the ethical safeguards described under Instrument and Procedures). Because participants were elementary school pupils recruited through the school rather than self selected community members, they are referred to throughout this manuscript as “students,” rather than as “speakers” or “learners” of Mandailing in a general community sense. All 25 students in the Grade 5 class participated in the product trial, the pretest and posttest, and the user response questionnaire. No separate feasibility testing sample was used. Table 3 summarizes participation across the study stages.

Instrument and Procedures

The development procedure followed the five ADDIE stages. Needs analysis data were gathered through interviews with four informants, one from each of four roles: a customary leader (tokoh *adat*), a cultural figure (*budayawan*), a classroom teacher, and a student, each interviewed on 4 May 2026 during the Analysis stage. Interviews with the customary leader and cultural figure indicated that terms such as *markobar* (customary speech making), *horja* (customary celebration), dalihan na tolu (kinship system), bagas godang (traditional house), and itak gurgur (traditional food) are reported to be increasingly unfamiliar to younger community members. A technology access check found that all students in the target class had access to an Android smartphone, though with limited stable internet access at home. This finding informed the decision to design the product as a lightweight, low data website that would remain usable on a limited or intermittent connection (see Offline and Access Claims in Development Stage Results). Observation was used to examine Mandailing-vocabulary use in students' everyday environment and cultural practices relevant to local wisdom, and to observe the product trial session for contextual data on ease of use. Semi structured interviews were conducted with customary leaders and local language teachers to gather Mandailing local wisdom content for the dictionary, and with a subset of respondents regarding their experience using the product.

Questionnaires were the main quantitative instrument, comprising (a) validation questionnaires completed separately by one material expert, one media expert, and one language expert, and (b) a user response questionnaire completed by 25 students after product use, covering perceived content, language, and media quality, ease of use, and perceived benefit. All questionnaires used a four point Likert scale (Strongly Agree = 4 to Strongly Disagree = 1). A Mandailing thematic vocabulary test was administered to the 25 participating students immediately before (pretest) and after (posttest) product use, to assess short term change in comprehension of the thematic vocabulary covered by the product. The test comprised 15 multiple choice items, administered within a 30 minute session, with a maximum possible score of 100. All 15 items were checked for validity, and all 15 were retained as valid. The pretest and posttest used the identical 15 items, presented in a different item order between the two administrations (for example, an item appearing as item 2 on the pretest could appear as item 10 on the posttest); this reduces the likelihood that students simply memorized answer positions, though content level practice or memory effects from encountering the same items twice within roughly an hour cannot be ruled out. Documentation included photographs, audio visual recordings of the research process, and written records of Mandailing local wisdom (e.g., customary texts and oral history notes) used as source material for the product's cultural content. School authorization for the study was obtained from SD UPTD 04 Simatahari. No formal ethics committee review or written parental/guardian consent was obtained; parents were informed of the study verbally rather than through a documented consent procedure, and no separate child assent process was recorded. Participation was voluntary, and student identities were coded (e.g., ID codes rather than real names) in the data analyzed for this study; pretest, posttest, and questionnaire results had no bearing on students' official school grades.

Data Analysis

Qualitative data from observation, interviews, and documentation were analyzed using an interactive technique comprising data reduction, data display, and conclusion drawing/verification, following standard qualitative analysis practice.
Quantitative data from the expert validation sheets and the user response questionnaire were analyzed descriptively: a total score was calculated for each assessment aspect, an average score was computed as the total score divided by the number of items and respondents, and the average was converted to a qualitative feasibility category following Widoyoko (2016), shown in Table 6. The product was considered feasible if both expert validation and user response averages fell in at least the "Feasible" category (average score > 2.51).

Pretest and posttest data from the 25 participating students were analyzed using the Normalized Gain (N Gain): N Gain = (Posttest Score minus Pretest Score) / (Maximum Score minus Pretest Score), where the maximum possible score is 100. Interpretation followed Hake's (1998) criteria in Table 7. Because the individual level N Gain category distribution reported below (0 High, 25 Moderate, 0 Low) requires a value per participant, N Gain was evidently computed for each of the 25 students individually and then averaged, rather than computed once from group means. Because this study used a one group design without a comparison condition, the descriptive pretest/posttest means and the N Gain index describe within group change only; they do not by themselves establish statistical significance or isolate the product's effect from alternative explanations such as repeated testing, maturation, or close alignment between test items and product content.

RESULT AND DISCUSSION

This section presents the findings of the ADDIE based development of the Thematic Digital Visual Dictionary for Mandailing, Indonesian. The product trial was conducted during the Implementation stage, 1 to 5 July 2026, with the 25 student Grade 5 class.

Analysis Stage Results

Needs analysis findings indicated that many respondents had limited understanding of thematic Mandailing vocabulary related to customary ceremonies, kinship, traditional cuisine, and traditional house architecture; interview data from the four informants (customary leader, cultural figure, teacher, and student; see Needs Analysis above) identified several culturally significant but reportedly under recognized terms. All 25 students in the target class were confirmed to own an Android smartphone, with limited stable home internet access, informing the decision to keep the product lightweight and low data rather than dependent on a fast connection.

Design Stage Results

Based on the needs analysis, five thematic vocabulary categories were designed at this stage: (1) customs and traditional ceremonies, (2) kinship systems and forms of address (dalihan na tolu), (3) traditional cuisine, (4) architecture and traditional houses, and (5) arts and traditional attire. As described in Development Stage Results below, a sixth category, cultural symbols, was added during product development.

Development Stage Results

The final product is a Thematic Digital Visual Dictionary for Mandailing, Indonesian, comprising six thematic vocabulary categories, forms of address (*sapaan*), customs (*adat*), food (*makanan*), house (*rumah*), clothing (*pakaian*), and cultural symbols (*simbol*), together totaling 105 vocabulary entries, a word search feature, illustrations for each entry, native speaker pronunciation audio, and a self assessment quiz feature. The final entry count of 105 reflects a narrowing of the initial planning estimate of "approximately 250" entries stated at the Design stage; Following expert validation, the draft 1 prototype was revised into a final draft 2 product through corrections to the visual display and to spelling/orthography; no content or structural changes were made at this step. Once development was complete, the product was validated by two validators covering three domains: the material and language domains were both rated by Shopiah Anggraini Rambe, M.Hum., a lecturer at Universitas Al Washliyah (UNIVA) and a native speaker of Mandailing; and the media domain was rated by Selamat Subagio, M.Kom., also a lecturer at UNIVA. Because Shopiah Anggraini Rambe is a co author of this study, the material and language validation domains were not independent of the development team; this is disclosed here and treated as a limitation of the validation process. The media validator is not a co author and is, on present information, independent.

Product Specification

Table 2. Product Specification

Attribute	**Specification**
Platform / access route	Web based application, accessed by browser at kamus-bahasa-mandailing.page.gd (PHP based pages, e.g., kamus.php), with a separate admin login for content management.
Development software	Built as a standard PHP based web application.

Attribute	**Specification**
Online/offline mechanism	Confirmed: a standard browser accessed website (not an installable, offline caching app); described as lightweight and low data, suited to limited or intermittent connections, rather than offline capable.
Device / OS requirements	No special device or operating system requirement; usable on any standard, modern web browser.
Navigation and search	Two tier browsing structure confirmed from product screenshots: a curated "Sering Digunakan Sehari hari" (Frequently Used Daily) subset on the entry list page, plus the full "Semua Kosakata (A, Z)" (All Vocabulary, A, Z) alphabetical listing; a search bar ("*Cari kata*," with placeholder example "*ayah*") and an alphabet filter ("*Filter Abjad*") button are both visible on the browsing screen.
Quiz behavior	Self assessment quiz feature comprising 15 multiple choice questions.
Audio format	*[AUTHORS: audio file format (e.g., MP3) and approximate duration per entry still needed.]*
Number of entries	105 vocabulary entries
Thematic categories (final, six)	Forms of address (*sapaan*), customs (*adat*), food (*makanan*), house (*rumah*), clothing (*pakaian*), cultural symbols (*simbol*)
Accessibility features	*[AUTHORS: e.g., text size adjustment, captions/alt text for images, audio replay controls, please describe what is available]*

Content and Local Wisdom Curation

The 105 vocabulary entries were developed from needs analysis interviews with four informants (a customary leader, a cultural figure, a teacher, and a student; see Needs Analysis) and validated by the material and language experts, both native Mandailing speakers. Table 3 lists ten representative entries observed directly in the product's browsing screens, spanning all six thematic categories, to let readers see the product's actual content.

Table 3. Representative Vocabulary Entries

Category	**Indonesian Term**	**Mandailing Term**
Forms of address (*sapaan*)	*Ayah*	*Amang*
Forms of address (*sapaan*)	*Ibu*	*Umak*
Customs (*adat*)	*Musyawarah Adat*	*Martahi*
Customs (*adat*)	*Pernikahan Adat*	*Horja Godang*
Food (*makanan*)	*Sambal Tuk Tuk*	*Sambal Tuk Tuk*
Food (*makanan*)	*Lemang*	*Lomang*
House (*rumah*)	*Rumah Adat*	*Bagas Godang*
Clothing (*pakaian*)	*Ulos*	*Ulos*
Clothing (*pakaian*)	*Mahkota Adat*	*Bulang*
Symbols (*simbol*)	*Melambangkan Fungsi Raja Dan Namora Natoras*	*Burangir*

Audio and Visual Assets

A detail page screenshot (for the entry "*Adik*"/"*Anggi*") confirms that individual entries display a larger illustration, the Indonesian term and its Mandailing equivalent under clearly labeled headings ("*Bahasa Indonesia*" / "Bahasa Mandailing"), an audio playback control next to the Mandailing term, and a "Contoh Penggunaan" (Example Usage) section intended to show a sample sentence. For this entry, the example usage section read "Belum ada contoh kalimat" ("No example sentence yet"), indicating that example sentences, advertised on the product's landing page ("learn example sentence usage"), are not yet populated for at least some entries

Material Expert Validation Results

Table 4. Material Expert Validation Results

No	Assessment Aspect	Average Score	Category
1	Conformity of content with Mandailing local wisdom	3.40	Highly Feasible
2	Accuracy of thematic vocabulary selection	3.60	Highly Feasible
3	Depth and breadth of material	3.20	Feasible
4	Relevance of material to the literacy improvement goal	3.80	Highly Feasible
5	Currency and contextuality of material	3.40	Highly Feasible
	Overall Average	3.48	Highly Feasible

Media Expert Validation Results

Table 5. Media Expert Validation Results

No	Assessment Aspect	Average Score	Category
1	Quality of visual display and layout	3.60	Highly Feasible
2	Quality of supporting illustrations/images	3.40	Highly Feasible
3	Quality of pronunciation audio	3.80	Highly Feasible
4	Ease of application navigation	3.60	Highly Feasible
5	Feature interactivity (search and quiz)	3.40	Highly Feasible
	Overall Average	3.56	Highly Feasible

Language Expert Validation Results

Table 6. Language Expert Validation Results

No	Assessment Aspect	Average Score	Category
1	Accuracy of Mandailing, Indonesian spelling	3.50	Highly Feasible
2	Consistency of term usage	3.25	Feasible
3	Clarity of vocabulary meaning	3.75	Highly Feasible
4	Conformity with linguistic rules	3.50	Highly Feasible
	Overall Average	3.50	Highly Feasible

Validation results indicate the product meets feasibility criteria on the material, media, and language dimensions, each rated Highly Feasible by its respective single expert. On the material dimension, the highest rated indicator was relevance of the material to the vocabulary learning goal (3.80). On the media dimension, the highest rated indicator was pronunciation audio quality (3.80). On the language dimension, the highest rated indicator was clarity of vocabulary meaning (3.75), while consistency of term usage scored lower, in the Feasible category (3.25), a note for continued attention to spelling and terminology consistency given Mandailing's regional dialectal variation.

The single validator per domain design means each “Overall Average” above is a mean across that validator's own rating items, not a measure of agreement between raters. As noted in Development Stage Results, the material and language domains were both rated by the same validator, Shopiah Anggraini Rambe, M.Hum. (a native Mandailing-speaking lecturer at UNIVA and a co author of this study), while the media domain was rated independently by Selamat Subagio, M.Kom. (UNIVA). Because one validator is a co author, the material and language scores in Tables 10 and 12 should be read as a non independent assessment, not as external expert consensus; this is treated as a limitation (see Conclusion and Limitations). The combined score of 3.51 (Highly Feasible) is the unweighted mean of the three domain overall averages, (3.48 + 3.56 + 3.50) / 3 = 3.51, despite the three domains using different numbers of rating items (5, 5, and 4 respectively); this is noted as a simplification rather than a weighted synthesis.

Implementation Stage Results

The trial was conducted during 1 to 5 July 2026 with the 25 participating Grade 5 students (15 female, 10 male; age 10 to 11 years) at SD UPTD 04 Simatahari. The session began with the pretest, followed by approximately 60 minutes of

independent product use, then the posttest, and finally the user response questionnaire. Because the posttest immediately followed the pretest and the product use session, and used the same 15 items (in a different order) as the pretest, the observed score change may partly reflect item content familiarity or practice effects, in addition to any learning attributable to the product; this is discussed further under Discussion and Limitations.

User Response Questionnaire Results

Table 7. User Response Questionnaire Results

No	Assessment Aspect	Average Score	Category
1	Content Feasibility	3.42	Highly Feasible
2	Language Feasibility	3.35	Highly Feasible
3	Media Feasibility	3.56	Highly Feasible
4	Ease of Use	3.48	Highly Feasible
5	Benefit	3.61	Highly Feasible
	Overall Average	3.48	Highly Feasible

Respondents rated the product positively overall, with the highest rated aspect being perceived benefit (3.61) and the lowest being perceived language quality (3.35), still within the Highly Feasible category. Overall Average = 3.48. Table 13 currently reports domain means only.

Pretest and Posttest Results

The pretest and posttest data from the 25 respondents are presented in the table below.

Table 8. Pretest and Posttest Results

Statistic	Pretest	Posttest
Mean	54.12	79.40
Standard Deviation (SD)	5.68	4.67
Minimum Score	45	72
Maximum Score	65	88

Mean scores rose from 54.12 (pretest) to 79.40 (posttest), an average increase of 25.28 points. The average N Gain was 0.56 (Moderate category; Hake, 1998); by the individual N Gain distribution as reported, 0 of 25 respondents (0%) fell in the High category, 25 (100%) in the Moderate category, and 0 (0%) in the Low category. This pattern is consistent with, but does not on its own establish, a benefit specific to the local wisdom based thematic approach relative to other vocabulary learning methods; alternative explanations are discussed below

Evaluation Stage Results

Formative evaluation was carried out throughout development, reflected in the expert validation results and accompanying revisions. Based on the data collected: (1) the product was rated Highly Feasible by three experts, one per domain, with a combined average of 3.51; (2) the product was rated positively for practicality by 25 users, with an average of 3.48; and (3) mean Mandailing thematic vocabulary scores increased from pretest to posttest, with an average N Gain of 0.56 (Moderate category). Because this evaluation used a single group, an immediate posttest, and no comparison condition, point (3) is best described as a preliminary, short term association between product use and thematic vocabulary score change, not as demonstrated effectiveness in a causal or generalizable sense.

The feasibility findings are broadly consistent with Rahmawati and Suliyatun (2021) and Haryani and Yustha (2021), both of which reported Highly Feasible ratings for locally oriented dictionary media developed through Research and Development approaches. Unlike those two studies, and responding to a gap noted in Bogar et al. (2025), this study also collected pretest, posttest data on thematic vocabulary scores rather than stopping at product feasibility. The observed increase in mean thematic vocabulary scores (N Gain = 0.56, Moderate category; Hake, 1998) is consistent with, but does not by itself demonstrate, a causal benefit of the local wisdom based thematic approach specifically. Because the design used one group, one brief exposure, an immediate posttest with the same 15 items (only reordered) as the pretest, and no comparison condition, several alternative explanations cannot be ruled out: direct item content memory from encountering the identical 15 items roughly an hour earlier; the novelty of using a new digital tool; facilitator or teacher support during the session; and ordinary within session familiarization with the tested content regardless of medium. Distinguishing these explanations from a genuine, retained learning benefit would require, at minimum, a comparison condition (for example, a class using a conventional printed dictionary or no supplementary material), parallel (non identical) pretest and posttest forms, a delayed posttest to assess retention, and ideally a transfer or applied use measure.

We therefore treat the score increase as preliminary evidence that the product is usable and engaging enough to be associated with a short term thematic vocabulary gain, not as evidence that it improves general language literacy, strengthens cultural identity, or revitalizes the Mandailing language. Those broader outcomes were not measured in this study; they remain motivating premises, consistent with the Society 5.0 framing (Fukuyama, 2018), for future research designed specifically to test them. The lowest rated indicator across all three expert validation domains was consistency of term usage on the language dimension (3.25, still Feasible; see Table 12), reflecting the dialectal variation across Mandailing-speaking regions noted by the language expert; this is a concrete, actionable area for revision, for example, adopting and documenting a single spelling convention (see Content and Local Wisdom Curation), before any wider release. For teachers, the preliminary results

suggest the product may be usable as a supplementary vocabulary introduction tool within a single class period, provided a teacher is present to support use (see Implementation Fidelity); it is not, on this evidence, a substitute for sustained classroom instruction. For community and customary authority partners, the product's dependence on a small number of informants and validators (see Content and Local Wisdom Curation) means broader community review, beyond the four needs analysis informants and three experts consulted here, would strengthen confidence in the product's cultural accuracy before it is presented as representing the Mandailing community's local wisdom more generally.

CONCLUSION

This study developed and conducted a preliminary evaluation of a Thematic Digital Visual Dictionary for Mandailing, Indonesian based on local wisdom, using the ADDIE model. Three experts, one material, one media, one language, rated the product Highly Feasible (combined mean = 3.51 on a 4 point scale), and 25 elementary school users rated it positively for practicality and usefulness (mean = 3.48). In a one group, one session pretest, posttest trial with the same 25 students, mean Mandailing thematic vocabulary scores rose from 54.12 to 79.40, with an average N Gain of 0.56 (Moderate category). Within the bounds of this preliminary, single group design, these findings support the product's feasibility and short term usability as a thematic vocabulary learning tool for elementary school students. They do not, on their own, establish improvement in general language literacy.

REFERENCES

Aprilia, P., Amri, K., Annahari, A., Parapat, A. A., Hasibuan, S. R., & Harahap, R. A. (2025). Mandailing language maintenance in the era of globalization: Challenges, strategies, and prospects. *Journal Angkola Mandailing, 1*(1), 25 to 37. https://doi.org/10.64733/journalangkola-mandailing.v1i1.99

Arikunto, S. (2022). *Prosedur penelitian: Suatu pendekatan praktik* [Research procedures: A practical approach]. Rineka Cipta.

Ayatrohaedi. (1986). *Kepribadian budaya bangsa (local genius)* [National cultural personality (local genius)]. Pustaka Jaya.

Badan Pengembangan dan Pembinaan Bahasa. (2024). *Status bahasa daerah di Indonesia* [The status of local languages in Indonesia]. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.

Bogar, D. S., Raymond, C. P., Yeninar, R. F., & Dharsono, W. W. (2025). Pengembangan aplikasi kamus elektronik bahasa Napan berbasis Android untuk pelestarian bahasa daerah di Papua Tengah [Development of an Android based Napan language electronic dictionary application for local language preservation in Central Papua]. *Jurnal Ilmiah Sistem Informasi dan Teknik Informatika (JISTI), 8*(1), 54 to 64. https://doi.org/10.57093/jisti.v8i1.258

Branch, R. M. (2009). *Instructional design: The ADDIE approach*. Springer.

Fauzi, H. N., & Rosliyah, Y. (2020). Persepsi pembelajar terhadap media kamus tematik berbasis web bagi pembelajar kosakata bahasa Jepang tingkat dasar [Learners' perceptions of a web based thematic dictionary medium for beginner Japanese vocabulary learners]. *Jurnal Pendidikan Bahasa Jepang.*

Fukuyama, M. (2018). Society 5.0: Aiming for a new human centered society. *Japan Spotlight, 27*, 47 to 50.

Hake, R. R. (1998). Interactive engagement versus traditional methods: A six thousand student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64 to 74. https://doi.org/10.1119/1.18809

Haryani, T., & Yustha, Y. (2021). Pengembangan kamus bergambar anak dwibahasa (Dayak Ma'anyan Inggris) dalam pembelajaran muatan lokal [Development of a bilingual children's illustrated dictionary (Dayak Ma'anyan English) for local content learning]. *Anterior Jurnal, 20*(2), 141 to 152. https://doi.org/10.33084/anterior.v20i2.2106

OECD. (2023). PISA 2022 results: Reading literacy. *OECD Publishing.*

Rahmawati, R. D., & Suliyatun, K. (2021). Pengembangan kamus saku Arab Indonesia untuk meningkatkan keterampilan berbicara bahasa Arab [Development of an Arabic Indonesian pocket dictionary to improve Arabic speaking skills]. *Jurnal Education and Development, 9*(3), 196 to 199.

Sianturi, R. A., Simanjuntak, S., Simanjuntak, M., & Purba, G. (2019). Applying user centered design to mobile application for children. *IOP Conference Series: Journal of Physics: Conference Series, 1175*, 012094. https://doi.org/10.1088/1742-6596/1175/1/012094

Sugiyono. (2021). *Metode penelitian kuantitatif, kualitatif, dan R&D* [Quantitative, qualitative, and R&D research methods]. Alfabeta.

Tambaip, B., Wayangkau, I. H., Suwarjono, S., Loupatty, M., Adam, A. F., & Hariyanto, H. (2020). Preservation of the Muyu Indigenous Language with an Android based Dictionary. *Emerging Science Journal, 4*, 85 to 101. https://doi.org/10.28991/esj-2021-SP1-06

Tarp, S., & Gouws, R. H. (2023). A necessary redefinition of lexicography in the digital age: Glossography, dictionography and implications for the future. *Lexikos, 33*(1), 425 to 447.

Widoyoko, E. P. (2016). *Evaluasi program pembelajaran: Panduan praktis bagi pendidik dan calon pendidik* [Evaluation of learning programs: A practical guide for educators and pre service educators]. Pustaka Pelajar.