

Babelum: A Published Video Game for Second-Language Vocabulary Acquisition

Theory-informed interactive learning design and controlled comparative evaluation

Paulo Jorge Pereira Martins | Centre for Informatics and Systems, University of Coimbra (CISUC) | E-mail: paulopm@dei.uc.pt | CV: www.paulojorgepm.net/cv

1. Introduction and Goal

Babelum is a digital game for second-language vocabulary acquisition, released on Steam in April 2026 (App ID 4391760). It operationalizes the pedagogical concept proposed in Paulo Martins' 2013 master's thesis, translating Montessori-inspired lexical practice into a production system with a snake-style core gameplay loop that supports repeated retrieval under engagement-oriented interaction [3].

Digital game-based language learning (DGBLL) shows generally positive but heterogeneous effects; outcomes depend on intervention quality, task design, and evaluation rigor [1, 6–8].

Study End Goal: Quantify the incremental vocabulary-learning effect of Babelum versus matched Montessori-style flashcards under controlled time-on-task.



Figure 1: Babelum gameplay with lexical interaction and retrieval tasks.

2. Background and Thesis Continuity

The original thesis defined a *ludic-pedagogical* prototype for L2 lexical initiation centered on concrete lexical items and active learner manipulation [3]. Babelum keeps this pedagogical spine but operationalizes it as a deployable multilingual game artifact with measurable gameplay telemetry.

Retained design principles:

- Vocabulary-first progression before grammar-heavy instruction.
- Multimodal lexical cues (object/image, orthography, and audio).
- Retrieval-oriented loop (perception, recall, and letter-order confirmation).
- Progressive challenge/reward to sustain persistence.
- Low-friction interaction patterns to reduce onboarding cost.

3. Knowledge Gap and Research Questions

Recent reviews report methodological weaknesses in DGBLL studies: limited active controls, small sample sizes, and weak delayed-retention designs [6–9].

Gap statement: Babelum is deployed publicly, but controlled evidence for incremental effect size versus a matched flashcard baseline is still missing.

RQ1: Under matched time-on-task, does Babelum improve immediate lexical recall versus Montessori-style flashcards?

RQ2: Does Babelum improve listening-based recognition accuracy and response speed?

Gameplay Modes and Learning Targets



Figure 3: Regular Levels - Lexical Acquisition.



Figure 4: Runner Levels - Phonological Awareness.



Figure 5: Story Levels - Verb and Adjective Learning.

RQ3: Are gains preserved in a delayed retention window (24–48h)?

RQ4: Which gameplay mechanisms most strongly predict gain (retrieval cycles, multimodal cueing, reward pacing)?

4. Methodology (Learning Design)

The implementation method maps language-learning theory to concrete mechanics so pedagogical intent is preserved during interaction design and progression tuning.

- DGBLL and gamification evidence motivates mechanism-aware design and rigorous evaluation [1, 6–9].
- Involvement load principles guide active lexical processing through goal-oriented retrieval tasks [2].
- Retrieval practice with repeated encounters is used to strengthen retention [5].
- Multimedia-learning principles align visual, orthographic, and audio channels during each lexical event [4].
- Montessori-compatible learner agency is implemented via direct manipulation within structured goals.

5. Construct Model

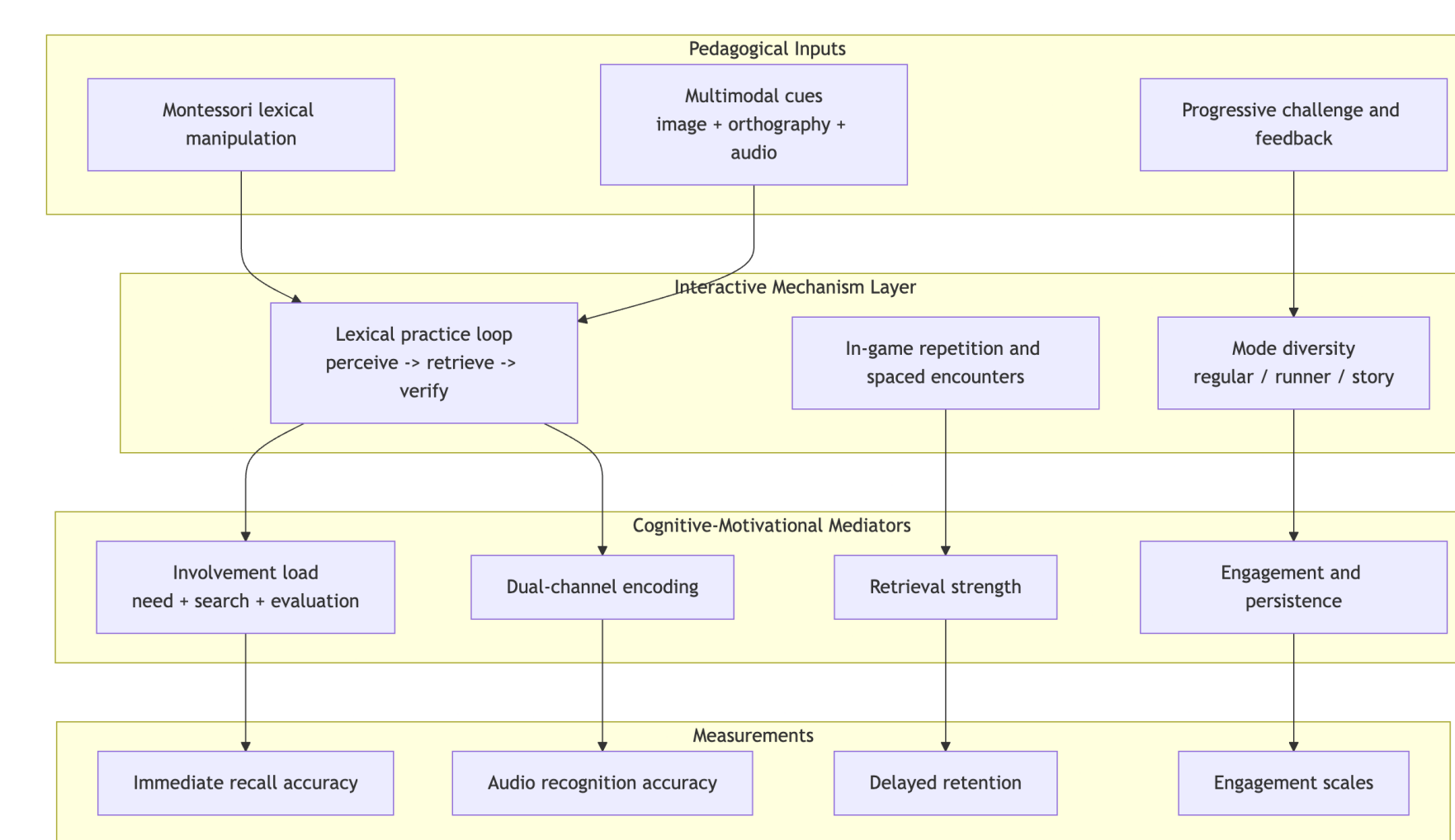


Figure 2: Construct model linking multimodal lexical exposure and retrieval interaction to immediate recall, recognition, and delayed retention outcomes.

6. Current Artifact Status (Repository Evidence)

| Dimension | Current measurable status |
|-------------------|--|
| Level Corpus | 30 authored levels across Farm, City, and House domains (10 each). |
| Mode Portfolio | Four playable modes: Regular (lexical acquisition), Exploration (practice and consolidation), Runner (phonological training), and Story (contextual verbs/adjectives). |
| Lexical Targets | Primary lexical classes include nouns, verbs, and adjectives. |
| Language Coverage | Ten target language packs: EN, PT, ES, FR, DE, IT, JA, KR, CN, and IN. |
| Instrumentation | Telemetry captures score, elapsed time, steps, stars, and lexical completion signals. |
| Deployment Status | Public Steam release in April 2026 (App ID 4391760). |

7. Comparative Study and Next Steps

Next step: execute a controlled study comparing Montessori-style flashcards versus Babelum gameplay under matched

time-on-task, with immediate and delayed post-tests. This protocol is the immediate next research phase; participant recruitment is active and data collection will start after pre-registration closure.

Design: randomized, counterbalanced within-subject crossover (AB/BA), with matched lexical sets and matched exposure duration.

Participants:

- Adult volunteers (target: university/community L2 learners).
- Baseline profiling: L1/L2 background, prior game habits, self-rated proficiency.

Procedure:

1. Consent and baseline instruments.
2. Condition A (flashcards, set X) and Condition B (Babelum, set Y), counterbalanced.
3. Immediate post-condition tests: spelling/reading recall and audio recognition.
4. Short break and crossover.
5. Delayed retention test.

Outcome variables:

- Primary: proportion of correctly recalled target words.
- Secondary: listening recognition accuracy, response latency, engagement/workload scales, and error taxonomy.

Analysis plan:

- Mixed-effects regression (condition/order/proficiency fixed effects; participant/item random effects).
- Paired effect sizes with confidence intervals.
- Robustness checks for sequence and novelty effects.

8. Validity, Ethics, and Reproducibility

- Counterbalancing and lexical-set matching reduce order and difficulty bias.
- Standardized audio and instructions reduce instrumentation variance.
- Pre-registered recruitment and protocol transparency improve reproducibility.
- Data governance: consent, anonymization, and participant withdrawal rights before analysis lock.

9. Open Questions

- How stable are effects across proficiency bands and age groups?
- Do specific mechanics (audio cueing, retrieval pacing) explain differential gains?
- Do mode-specific patterns (regular, runner, story) transfer to delayed retention?

10. Conclusion

Babelum is a production-ready multilingual artifact grounded in language-learning theory. The planned controlled evaluation targets reproducible effect estimation against an active baseline, advancing evidence on mechanism-level learning gains in DGBLL.

Want to participate in the study? Pre-register via:
<https://forms.gle/T6TVASBzLkFnEn6D6>

References

- [1] David B. Clark, Emily E. Tanner-Smith, and Sarah S. Killingsworth. 2016. Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis. *Review of Educational Research* 86, 1 (2016), 79–122. doi:10.3102/0034654315582065
- [2] Jan H. Hulstijn and Baita Laufer. 2001. Some Empirical Evidence for the Involvement Load Hypothesis in Vocabulary Acquisition. *Language Learning* 51, 3 (2001), 539–558. doi:10.1111/0023-8333.00164
- [3] Paulo Jorge Pereira Martins. 2013. *Humanidades Interativas: Problematização do Processo de Desenho e Construção de um Videjogo Lúdico-Pedagógico*. Master's Thesis, Universidade do Minho.
- [4] Richard E. Mayer. 2005. Cognitive Theory of Multimedia Learning. In *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press, Cambridge, UK, 31–48. doi:10.1017/CBO9780511816819.004
- [5] Henry L. Roediger and Jeffrey D. Karpicke. 2006. Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. *Psychological Science* 17, 3 (2006), 245–255. doi:10.1111/j.1467-9280.2006.01693.x
- [6] Michael Sailer and Lisa Homner. 2020. The Gamification of Learning: A Meta-analysis. *Educational Psychology Review* 32, 1 (2020), 77–112. doi:10.1007/s10648-019-09498-w
- [7] Yu-Ling Tsai and Chin-Chung Tsai. 2018. Digital Game-Based Second-Language Vocabulary Learning and Conditions of Research Designs: A Meta-Analysis Study. *Computers & Education* 125 (2018), 345–357. doi:10.1016/j.compedu.2018.06.020
- [8] Lu Yang and Rui Li. 2024. Contextualized Game-Based Language Learning: Retrospect and Prospect. *Journal of Educational Computing Research* 62, 1 (2024), 137–155. doi:10.1177/07356312231189292
- [9] Ruofei Zhang, Di Zou, and Gary Cheng. 2023. Learner Engagement in Digital Game-Based Vocabulary Learning and Its Effects on EFL Vocabulary Development. *System* 119 (2023), 103173. doi:10.1016/j.system.2023.103173

Babelum's Steam Page



<https://store.steampowered.com/app/4391760>

Study Registration



<https://forms.gle/T6TVASBzLkFnEn6D6>

Copyright © Retro Tales | 2026 | <https://retrotales.eu/>

