Conference Name: 2025 EdTec – International Conference on Education & Learning Technology, 08-09 September, Rome
Conference Dates: 08-Sep- 2025 to 09-Sep- 2025
Conference Venue: Courtyard by Marriott Rome Central Park, Via Moscati 7 •00168 Rome • Italy
Appears in: PUPIL: International Journal of Teaching, Education and Learning (ISSN 2457-0648)
Publication year: 2025

Ezekiel & Olagunju, 2025

Volume 2025, pp. 519-552

DOI- https://doi.org/10.20319/ictel.2025.519552

This paper can be cited as: Ezekiel, D.O. & Olagunju, E.(2025). Reimagining Language Learning through Artificial Intelligence and Gamification for Globally Mobile Learners in Non-Immersive Contexts (with a Focus on DaZ/DaF). 2025 EdTec – International Conference on Education & Learning Technology, 08-09 September, Rome. Proceedings of Teaching & Education Research Association (TERA), 2025, 519-552

REIMAGINING LANGUAGE LEARNING THROUGH ARTIFICIAL INTELLIGENCE AND GAMIFICATION FOR GLOBALLY MOBILE LEARNERS IN NON-IMMERSIVE CONTEXTS (WITH A FOCUS ON DAZ/DAF)

Dorcas O. Ezekiel
Freiburg University of Education, Freiburg, Germany
dorcasezekielofficial@gmail.com

Ezekiel Olagunju
Department of Foreign Languages, Obafemi Awolowo University, Ile-Ile, Osun State. Nigeria
olagunju34@oauife.edu.ng

Abstract

This article advances a critical reimagining of German language acquisition for globally mobile learners navigating non-immersive, digitally mediated contexts. Positioned at the intersection of Second Language Acquisition (SLA), cognitive linguistics, and human-computer interaction, the study interrogates the structural and pedagogical limitations of current Mobile-Assisted Language Learning (MALL) platforms—particularly their failure to support integrated, skill-balanced development in linguistically demanding contexts such as DaZ/DaF. Drawing on Design-Based

Research (DBR), the authors propose a multi-tiered mobile learning ecology anchored in a prototypical application that combines AI-driven adaptivity with narrative-based gamification. This hybrid architecture simulates immersive conditions through city-based tasks, interactive dialogue flows, and CEFR-aligned communicative scenarios. Empirical findings from a purposively sampled cohort of fifty globally mobile learners reveal persistent cognitive and affective challenges, including fragmented input, insufficient feedback, and metacognitive disorientation. The prototype responds by embedding linguistic scaffolding, affect-sensitive interaction design, and dynamic personalization mechanisms to bridge the gap between learner autonomy and structured progression. By repositioning non-immersed learners as central actors in the design of pedagogically rigorous, technologically augmented language learning environments, this study articulates a scalable and theoretically grounded framework for post-immersion language pedagogy. In so doing, it offers a substantive contribution to ongoing debates in SLA, educational technology, and transnational language instruction—proposing not merely an app, but a paradigm shift in how language acquisition is imagined and operationalized in the 21st century.

Keywords:

Simulated Immersion, AI-Supported Language Learning, Gamified Mobile Pedagogy, Non-Immersive German Acquisition, Transnational DaF/DaZ Learners, Design-Based Research in Language Education, Adaptive Conversational AI, Cognitive–Affective Scaffolding, Task-Based Mobile Learning, Digital Learning Ecologies

1. Introduction

Globally mobile learners constitute a rapidly expanding and heterogeneous demographic, comprising individuals who relocate frequently for professional, academic, or personal reasons and who must often acquire new languages to navigate these transnational realities. Unlike traditional learners situated in immersive contexts, these individuals face distinct constraints: geographic displacement, irregular access to native speakers, and limited integration into the socio-cultural fabric of the target language environment. Consequently, they rely heavily on non-immersive learning contexts, where instruction is often fragmented, asynchronous, and technologically mediated. For learners of German - a language of considerable relevance in global education, research, diplomacy, and industry - these conditions present particularly formidable challenges. German's linguistic architecture is notoriously complex, characterized by a demanding four-case grammatical system, strict rules for verb placement, and the frequent use of lengthy compound nouns (e.g., *Donaudampfschifffahrtsgesellschaftskapitän*). Furthermore, the language's distinctive phonology, including umlauted vowels and consonantal sounds uncommon in many other languages, compounds the difficulty for non-native speakers and learners without sustained, real-world interaction. These structural and phonetic demands underscore the urgency of developing pedagogical models that transcend physical immersion and meaningfully address the cognitive, affective, and practical needs of this unique learner group.

The advent and widespread adoption of Mobile-Assisted Language Learning (MALL) technologies have dramatically transformed the landscape of language education, offering learners an unprecedented degree of autonomy, flexibility, and accessibility. By enabling engagement with language materials across temporal and spatial boundaries, these tools have ostensibly lowered traditional barriers to instruction, particularly for geographically dispersed or time-constrained individuals. Empirical data confirm that the vast majority of language learners report prior use of mobile applications and digital platforms as integral components of their learning process. However, a significant pedagogical gap remains: despite their accessibility and popularity, existing non-immersive technologies often fall short of supporting the deep, sustained engagement required for authentic language acquisition, particularly in a linguistically demanding language like German.

Learners frequently cite persistent difficulties in cultivating spontaneous communicative competence, internalizing grammatical precision, and maintaining long-term motivation without the embodied, contextual reinforcement that immersion provides. Moreover, many current applications lack the sophistication to deliver dynamic, personalized feedback or simulate complex real-world language use. This disconnect is further exacerbated by psychological and generational barriers: while younger users may gravitate toward gamified platforms, older learners often express ambivalence or outright resistance, shaped by factors such as perceived digital competence and openness to innovation. These disparities underscore a critical misalignment between the functional design of current MALL tools and the cognitive, emotional, and pragmatic needs of diverse learner populations. What emerges is a clear imperative to reimagine mobile language learning, especially for structurally complex languages like German - not merely as a matter of convenience, but as a domain requiring thoughtful pedagogical engineering, adaptive personalization, and immersive simulation.

This report explores how the strategic application of advanced AI and sophisticated gamification within a mobile application can address these limitations, effectively reimagining German language learning. The proposed solution aims to create a "pseudo-immersive" environment that fosters deep engagement, provides adaptive feedback, and supports holistic skill development across all four language skills, thereby bridging the experiential gap inherent in non-immersive contexts. The subsequent sections will delve into the current landscape of German language learning, analyze existing technological tools, examine the linguistic and cognitive dimensions of German acquisition, explore relevant theoretical frameworks, and finally, propose an innovative design for an AI-gamified mobile application, supported by research findings and user feedback.

2. Literature Review

The field of language learning has undergone a radical transformation over the past two decades, particularly with the rise of Mobile-Assisted Language Learning (MALL), Artificial Intelligence (AI) -mediated instruction, and gamified environments. Yet, despite the proliferation of research in these individual domains, relatively few studies have synthesized them in relation to globally mobile learners operating in non-immersive contexts. This literature review maps out key strands of scholarship relevant to the present study, highlighting both theoretical foundations and empirical findings across four interlocking domains: (1) non-immersive language learning, (2) digital language pedagogy and MALL, (3) gamification in second language acquisition, and (4) AI-enhanced language learning.

2.1.1 Non-Immersive Language Learning: Constraints and Cognitive Burdens

Much of the foundational research in second language acquisition emphasizes the centrality of immersion and social interaction. Immersive environments provide real-time feedback, cultural context, and social pragmatics that foster linguistic fluency and communicative competence. However, for globally mobile learners - individuals who study or work across multiple locations- access to sustained immersive environments is often unattainable. These learners frequently rely on fragmented, asynchronous digital tools, which lack the intensity, feedback-rich environment, and spontaneity that immersion provides. Studies by Benson and Reinders (2011) have shown that self-directed, mobile learners often experience plateaus in fluency, particularly in productive skills like speaking and writing, unless scaffolding and interaction are robustly integrated into the learning environment.

2.1.2 Mobile-Assisted Language Learning (MALL): Promise and Limitations

MALL has emerged as a dominant paradigm in the digital transformation of language education, offering accessibility, flexibility, and modularized content delivery. Language apps such as Duolingo, Babbel, and Busuu exemplify this shift, employing multimedia, gamification, and adaptive algorithms to simulate learning experiences. Yet, while these platforms have demonstrated success in initial engagement and vocabulary acquisition, they often fail to develop linguistic competencies such as syntactic mastery, pragmatic nuance, and discursive cohesion skills, especially critical in German due to its grammatical complexity. Furthermore, research highlights a misalignment between app design and learner diversity: older adults and learners from non-Western

academic cultures frequently report alienation or low digital efficacy (Stockwell & Hubbard, 2013), suggesting that current MALL ecosystems inadequately serve heterogeneous, transnational learner populations.

2.1.3 Gamification: Motivation, Memory and Meaningful Learning

The use of game mechanics in educational environments has gained traction as a motivational tool, particularly in contexts where sustained learner engagement is a challenge. In second language acquisition, gamification has been shown to increase learner retention, reduce anxiety, and facilitate long-term memory encoding. However, critics caution against superficial implementation, warning that points, badges, and leaderboards may reward surface-level performance rather than deep linguistic processing (Thorne et al., 2015). Recent work calls for "meaningful gamification", which integrates narrative immersion, autonomy, and challenge elements more likely to simulate authentic communication and cognitive engagement. This research underscores the potential of city-based, narrative-driven, and context-specific gameplay as an alternative to reductive gamified modules.

2.1.4 Artificial Intelligence and Adaptive Language Instruction

AI's role in education has expanded from predictive analytics to dynamic, personalized instruction. In language learning, AI enables real-time feedback, adaptive content delivery, and natural language processing (NLP) for speaking and writing practice. Systems like Grammarly and speech-recognition tools used in apps like Elsa Speak exemplify the potential of AI to scaffold performance and reduce fossilized errors. However, the field remains in its infancy when it comes to integrating AI seamlessly with holistic skill development, particularly in under-researched areas like German syntax correction, case usage feedback, and pragmatic phrasing. Moreover, the AI models embedded in current language apps tend to be black-box systems, offering limited learner transparency and interaction with the feedback process (Godwin-Jones, 2018). There is, therefore, an urgent need for learner-facing, pedagogically grounded, and linguistically rigorous AI systems tailored to the structural challenges of German.

2.2 Synthesis and Gap

Despite parallel advancements in MALL, gamification, and AI-enhanced pedagogy, few studies have explored their combined potential to support learners who are transnational, non-immersed, and navigating complex languages like German. Existing tools frequently emphasize vocabulary acquisition or rudimentary grammar drills while neglecting the communicative, affective, and metacognitive dimensions of language learning. Moreover, little empirical attention has been paid to the user experience of globally mobile learners, whose learning trajectories are often interrupted, device-dependent, and self-directed. This article addresses that gap by proposing a research-based, pedagogically aligned, and technologically integrated mobile app design that supports full-spectrum language development - reading, writing, listening, and speaking - within a game-based, AI-supported environment tailored for non-immersive learners of German.

3. Methodology

3.1 Research Design

This study adopted a Design-Based Research (DBR) approach to explore how Artificial Intelligence (AI) and gamification could be integrated into a mobile learning application tailored for globally mobile learners of German operating in non-immersive contexts. DBR was selected because it allows iterative collaboration between theory and practice, supporting the co-design of learning solutions informed by empirical user research and existing theoretical frameworks. This approach focused on conceptualizing and refining key design features and pedagogical principles rather than building a fully functional application prototype. Instead, high-fidelity wireframes, screenshots, and interaction flows of the proposed app were developed and presented to participants to elicit detailed feedback. This user research informed the subsequent refinement and validation of the app design.

3.2 Participants and Sampling

To explore the realities of non-immersive German language acquisition and to guide the design of AI-supported, gamified mobile learning tools, the study employed a purposive sampling approach. Fifty adult learners of German were recruited, with ages ranging from 20 to 45. All participants were globally mobile individuals currently residing in non-German-speaking countries, thus excluded from immersion in native language environments. This criterion ensured that the data accurately reflected the everyday constraints and learning strategies of those studying German in

transnational, non-immersive contexts. Participants were carefully selected to reflect linguistic and cultural diversity. Their first languages (L1) included English, Yoruba, Igbo, Spanish, and Chinese, enabling the study to account for typological distance and cross-linguistic transfer challenges that may arise during German acquisition.

Participants were also chosen across a wide range of proficiency levels in German, from absolute beginners (A1) to highly advanced learners (C2), based on the Common European Framework of Reference for Languages (CEFR). This inclusion of the full CEFR spectrum enabled a layered analysis of how learners at different developmental stages engage with digital tools, what linguistic skills they struggle with most, and how app features such as AI or gamified tasks might scaffold different proficiencies in productive and receptive skills. Recruitment was conducted through online channels, including German language learning forums, Reddit communities, global Erasmus alumni groups, WhatsApp circles for international students, and targeted Facebook groups dedicated to language exchange.

To ensure that the participants could meaningfully engage with the app prototypes and provide relevant feedback, inclusion criteria were applied. Participants were required to have consistent access to a smartphone or tablet (either Android or iOS) and be willing to devote at least 30 minutes per day to German study. Importantly, participants were not enrolled in formal immersion programs and had not resided in a German-speaking country for more than six months, ensuring that the study remained focused on non-immersive learners without recent direct exposure to German-speaking environments. Exclusion criteria ruled out native German speakers as well as those with sustained immersion experience, in order to isolate the distinctive learning constraints and behaviors associated with non-immersion.

Data collection occurred in two interconnected phases, both qualitative in nature, in accordance with the methodological principles of Design-Based Research (DBR). The first phase involved semi-structured interviews and online focus groups designed to surface the lived experiences of non-immersive learners. These sessions, conducted via secure video conferencing platforms, explored participants' learning trajectories, motivational frameworks, frustrations, and tool usage patterns. Participants reflected on their use of well-known digital platforms such as Duolingo, Babbel, Memrise, Anki, and YouTube channels. Many spoke of monotony, the lack of productive language output, and the difficulty of achieving fluency or active vocabulary retention in the absence of social and linguistic immersion. These qualitative accounts revealed how cognitive and emotional strain - such as dealing with German morphosyntax or maintaining daily engagement

- often undercut learner persistence.

The second phase consisted of design feedback sessions focused on two experimental app concepts: *Kreuz und Quer*, a crossword-style vocabulary tool that emphasized morphological awareness and semantic categorization; and *Tandem Büro*, a swipe-based matching platform for real-time or simulated conversational exchanges, modelled loosely on Tinder but repurposed entirely for language practice. Participants were shown detailed wireframes, interactive user flows, and annotated screen mockups of both apps. These were accompanied by narrated walkthroughs to contextualize the pedagogical scaffolding and AI features embedded in each design. Crucially, participants were not only shown interface designs but were also walked through sample task flows (e.g. initiating a conversation match, responding to a gamified prompt, or navigating CEFR-leveled dialogues), enabling them to visualize how the app would function in everyday use.

After viewing the materials, participants were asked to provide feedback through structured questionnaires and open-ended discussions. These sessions captured critical user insights regarding visual design, feature usefulness, potential friction points, and their expectations for an AI-enhanced language companion. The feedback was particularly rich in relation to speaking and listening skills, which participants identified as the most underserved in current language learning apps. There was strong enthusiasm for simulated dialogues and AI-based fallback mechanisms, particularly where human interaction might be delayed or unavailable. Participants also responded positively to the idea of CEFR-informed progression models, but cautioned against overly rigid leveling that might penalize erratic learning paths common to adult learners.

All qualitative data (including interview transcripts, observational notes from feedback sessions, and written learner reflections) were subjected to thematic analysis. Coding proceeded inductively and iteratively, enabling analytic categories to emerge from within the data rather than being imposed externally. This approach was guided by the methodological sensibilities of grounded theory (Glaser and Strauss, 1967), particularly its commitment to theory generation through close, recursive engagement with participants' lived experiences. Thematic patterns were refined across multiple coding cycles and compared across learner profiles to ensure both internal coherence and interpretive depth. Emergent themes were then clustered into analytically distinct categories aligned with the core concerns of the study, namely, learner motivation, the constraints of non-immersive environments, autonomous learning strategies, and user interactions with the prototype. Categorization served not merely as an organizational device, but as a heuristic through which the data could be theorized in relation to broader questions of language learning design and

pedagogy.

Among the most prominent categories were learner strategies in the absence of immersion, such as self-dialogue, intentional media exposure, and code-switching to practice grammar patterns. A second theme concerned the limitations of mainstream tools—especially their lack of productive language challenges and low interactivity. A third set of categories focused on skill-specific difficulties: for instance, the challenge of parsing long compound nouns in reading, or producing grammatically complex spoken utterances under time constraints. Participants frequently reported difficulties with German morphosyntactic structures, such as the placement of modal verbs, separable prefixes, and article-case agreement, which often led to reduced confidence and communicative hesitation.

Another salient category concerned emotional and motivational regulation. Learners expressed that AI-driven scaffolding, if responsive and non-judgmental, could offer a form of psychological support, helping them sustain daily practice and regain confidence. Importantly, many participants welcomed gamification, but only if implemented in ways that respected adult learning sensibilities, avoiding cartoonish design or infantilizing rewards. Instead, they envisioned game-like experiences that simulate authentic environments, offer meaningful stakes (such as earning virtual currency), and encourage repeated engagement through socially or narratively embedded tasks.

The final analytic category focused on user expectations for AI integration. Participants expected the AI not just to correct errors, but to actively shape the learning experience—adapting task complexity, mimicking natural dialogue, simulating accents, and recognizing learner hesitation or difficulty mid-task. AI was framed as a bridge, not a replacement: a way to create continuity when human interaction was not available, especially in conversational practice, which many considered the most intimidating yet crucial skill.

These emergent categories served as both evaluative criteria and design parameters for refining the app concepts. Moreover, they informed the pedagogical rationale presented in subsequent chapters, including the need for immersive simulation within non-immersive settings, the ethical deployment of AI tutors, and the development of gamified environments that balance learner autonomy with structured skill-building.

4. Theoretical Framing

Findings and design implications were interpreted through multiple seminal lenses from Second Language Acquisition (SLA) theory, anchoring the research in established academic discourse:

- **Input Hypothesis (Krashen, 1985):** Emphasizes the importance of meaningful, comprehensible input slightly beyond learners' current proficiency ("i+1"). This informed the app's content scaffolding and AI-driven multimodal input delivery to optimize acquisition.
- **Output Hypothesis (Swain, 1985):** Highlights the necessity of active language production for developing fluency and accuracy. This guided the integration of AI-powered dialogue simulations and real-time feedback to encourage spontaneous spoken and written output.
- **Sociocultural Theory (Vygotsky, 1978; Lantolf, 2000):** Focuses on social mediation of learning through interaction, shaping the app's design to include peer collaboration, community gamified tasks, and learner agency within a virtual social environment.
- **Communicative Competence Model (Canale & Swain, 1980):** Outlines the multiple competences-grammatical, discourse, sociolinguistic, and strategic-that learners must develop. This influenced the balanced integration of these competences across speaking, listening, reading, and writing modules.
- **Motivational and Affective Theories (Deci & Ryan, 1985; Dörnyei, 2005):** Drawing on Self-Determination Theory and the L2 Motivational Self-System, these guided gamified reward systems, narrative identity features, and adaptive challenges to sustain motivation and reduce anxiety in non-immersive contexts.

5. Ethical Considerations

The study received institutional ethical approval. Participants provided informed consent and were assured confidentiality and anonymity. Participation was voluntary, with the right to withdraw at any point without consequence.

6. Results and Discussion

Globally mobile learners of German engage in language acquisition across fragmented, transient, and frequently unpredictable learning environments. These learners - often professionals, students, or migrants residing outside German-speaking countries - relocate regularly for academic, career, or personal reasons. As a result, they are structurally distanced from immersive settings and encounter logistical constraints that impede sustained exposure to the language. In this study, 50 participants aged 20-45 described their engagement with German as driven by long-term personal aspirations (e.g.: career migration, graduate study, or family integration), yet profoundly shaped by the daily realities of mobility, competing commitments, and separation from authentic German-speaking environments.

Learners consistently emphasized the *non-linearity* and *fluidity* of their study habits. Mobile-assisted learning enabled access to German in spontaneous moments - during commutes, late at night, between meetings - yet this flexibility also introduced a sense of pedagogical drift. Participants frequently described their routines as “piecemeal,” “looping,” or “all over the place.”

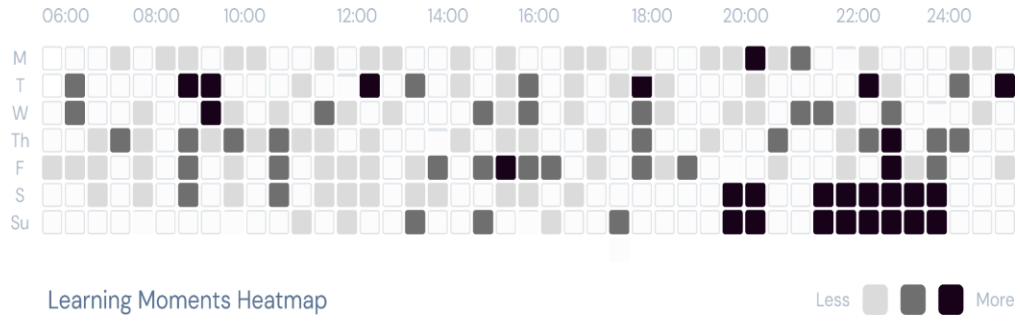

Figure 1: *Learning Moments Heatmap: A Heatmap Showing Common Times in which DAZ/DAF Learners Access Language Apps (E.G:, Late Evening, Commuting Hours, Weekends)*

Several participants reported regularly cycling between different apps and resources, often without a clear sense of cumulative progress. This form of digital multilingualism - what one participant called “language multitasking” - was efficient in frequency but deficient in pedagogical cohesion. Learners engaged with Duolingo exercises on their phones, watched German-language YouTube videos with English subtitles, or reviewed flashcards using spaced repetition software like Anki. While this form of digital multilingualism increased engagement, it lacked cohesion, scaffolding, and clear progression.

Thematic analysis of interviews and qualitative survey responses was conducted using *Kategorienbildung* (category formation), and initial codes such as “learning on the go,” “no grammar explanations,” “I just guess,” and “lots of apps, no guidance” were progressively clustered into higher-order analytic categories. These were then refined into the following analytic categories:

Table 1: *Key Analytic Categories Derived from the Research*

Analytic Category	**Description**
Fragmented Continuity	Frequent, short learning episodes without a pedagogical sequence. Language acquisition unfolded in short bursts without pedagogical sequencing, a phenomenon also observed in microlearning research (Kukulska-Hulme, 2020).
Analytic Category	**Description**
Unstructured Autonomy	Full learner control without guidance, corrective feedback, or curriculum structure.
Feedback Desert	Absence of corrective input, especially in speaking and writing, challenge particularly felt in speaking and writing tasks, where learners either avoid production entirely or practice without assurance of correctness.

Hollow Motivation	Engagement driven by gamified features rather than outcomes: users felt rewarded by streaks, badges, and scores, yet ultimately unsatisfied with their communicative progress.
Metacognitive Disorientation	Disconnection between effort and real-world communicative ability. One respondent captured this sentiment poignantly: *"I've studied German every day for six months and have the Duolingo diamond badge and a 200-day streak, but I still can't introduce myself properly in a real conversation."*

Participants routinely expressed frustration about not knowing whether they were improving. Many relied on digital metrics (streaks, XP points, levels) as proxies for competence, even as they doubted the validity of these metrics. This uncertainty affected self-efficacy and deepened what several described as "language anxiety", a cognitive state that led to avoidance of risk, especially in productive skills like speaking. The affective dimension of learning German in isolation was evident: learners expressed feelings of isolation, embarrassment when attempting to speak aloud, and uncertainty about how to gauge their own pronunciation or accuracy. The lack of interaction - what one respondent called "a silent classroom with no teacher," - created a flat learning experience devoid of dialogic exchange or responsive feedback.

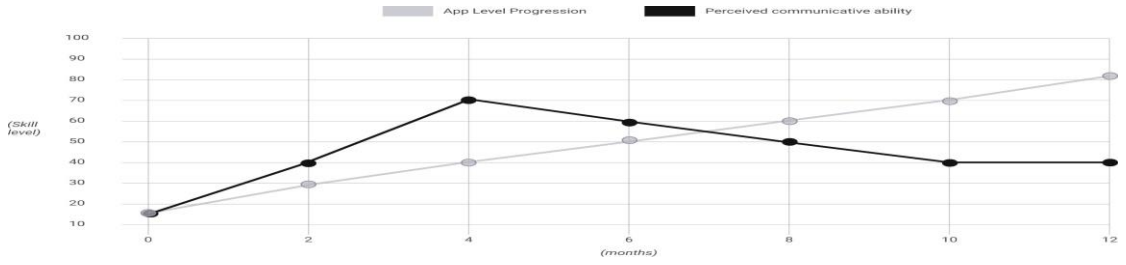

Figure 2: *Learner Confidence Vs. Platform Progression Chart: A Comparative Line Graph Showing Language Learners' Perceived Communicative Ability Versus App-Based Level Indicators*

Technologically, learners relied heavily on a core cluster of tools, with some variation by proficiency level. Duolingo was the most widely used app, especially among beginners (A1-A2), due to its gamified structure, low barrier to entry, and constant reinforcement. Intermediate and advanced learners (B1-C2), however, increasingly supplemented their app-based study with native content - German YouTube channels, Deutsche Welle videos, Reddit forums, or realia such as German recipes and podcasts. Nevertheless, the lack of comprehensible input and task-based structure rendered many of these resources cognitively overwhelming. For instance, learners described native content as either "way too fast" or "not relevant to what I need," highlighting a mismatch between linguistic difficulty and learner readiness.

Table 2: *Current Apps Used by Language Learners in Non-Immersive Contexts*

CEFR Level	**Primary Tools Used**	**Key Challenges Identified**
A1-A2	Duolingo, Babbel, YouTube Grammar	Grammar opacity, lack of speaking practice
B1-B2	Anki, Deutsche Welle, LingQ, Podcasts	Fluency plateau, tool fatigue, cognitive overload
CEFR Level	**Primary Tools Used**	**Key Challenges Identified**
C1-C2	Native texts, news, films, and Reddit forums	Pragmatic fluency, lack of natural interaction, register gaps

Despite the apparent variety of resources, participants across proficiency levels voiced dissatisfaction with the lack of integration between tools and skills. Tools were highly specialized, often targeting a single skill (e.g.: vocabulary memorization or listening practice) without facilitating holistic language development. Learners described feeling as though they were assembling a "Frankenstein's monster" of partial solutions, none of which supported sustained, interactive skill progression.

One Respondent summed this up by saying: *"Duolingo teaches words, Anki helps me remember them, YouTube explains grammar, but none of them help me speak or write properly. It's like I'm learning German in pieces."* This highlights a major conceptual limitation of many current tools: they are skill-specific but lack pedagogical orchestration. None of the platforms offered a unified environment for integrated development across all four core skills - speaking (Sprechen), listening (Hören), reading (Lesen), and writing (Schreiben).

At **A1-A2**, learners relied heavily on gamified apps like Duolingo, Babbel, and Busuu. These tools were perceived as approachable and motivating but limited in depth. Learners reported minimal grammatical understanding and frequent confusion about fundamental structures.

"I've completed the whole A1 tree on Duolingo and feel good when I see green lights and points, but then I freeze if I have to speak. I literally don't know what to say." - A1 learner, female, 29, Lagos-based DaZ/DaF Student

"I don't understand why 'der' becomes 'den,' even after 100 exercises. The app never tells me why." - A2 learner, female, 25, Nairobi-based self-taught DaZ/DaF learner

At the **B1-B2 level**, learners described entering a plateau. While vocabulary and comprehension improved, fluency and spontaneous sentence formation remained limited. Learners began supplementing with native materials, but struggled with the transition.

"You hit a wall around B2. You know lots of words, but you can't say what you want fast enough, and apps can't help you get there." - B2 learner, female, 34, Nigerian based DaZ/DaF Student

Key analytic codes for this group included "input mismatch," "fluency ceiling," and "self-curated overload." Learners at this stage created their own study ecosystems but lacked feedback mechanisms, structured progression, or guidance, resulting in cognitive fatigue and disorientation.

For **C1-C2 learners**, the challenge shifted from grammatical or lexical acquisition to *pragmatic and sociolinguistic fluency*. These learners could consume advanced materials and engage with German texts at a high level, but struggled with spontaneous speech, tone, idioms, and social register.

"At C1, I can read Kafka and follow politics, but I still hesitate before speaking because I'm not sure if I sound natural or robotic." - C1 learner, male, 38, Jakarta-based translator

"There's a ceiling you hit when you're not in Germany. You never get enough natural dialogue, sarcasm, real talk." - C2 learner, female, 41, New Delhi-based professor

The absence of real-time dialogic interaction and adaptive feedback emerged as a universal gap, intensifying with increased proficiency. Codes such as "Interactional Emptiness," "Expression without Calibration," and "Pragmatic Gaps" were prevalent among advanced learners.

Across all levels, participants reported difficulty integrating the four core language skills- **speaking (Sprechen), listening (Hören), reading (Lesen), and writing (Schreiben)** - in a cohesive manner. Tools emphasized isolated competencies, leading to asymmetrical proficiency. Learners were often stronger in receptive skills (listening and reading) than in productive ones (speaking and writing).

"I can write a decent paragraph and understand podcasts, but when it comes to talking to someone, I'm lost. There's no app that helps you practice real conversation without sounding like a robot." - B1 learner, female, 30, Kuala Lumpur-based NGO worker

Cognitive challenges were especially pronounced in morphosyntactic acquisition, a domain where learners frequently reported confusion. The case system, verb-second word order, separable-prefix verbs, and compound noun structures were cited as particularly difficult to grasp. Learners noted that app interfaces rarely paused to explain grammatical principles and that when explanations were available, they were too brief, oversimplified, or confusing. The lack of scaffolded grammar practice inhibited learners' ability to internalize rule-governed aspects of German syntax, a finding that echoes VanPatten's input processing theory (1996), which emphasizes the need for structured input and explicit attention to form in order to establish robust form-meaning mappings. Without adequate scaffolding, learners developed what might be called *surface fluency-comfortable* familiarity with linguistic forms, but limited productive flexibility.

Summarily, globally mobile learners of German in non-immersive contexts exhibit a form of **autonomous engagement without systemic coherence**. Their language journeys are marked by initiative and resourcefulness but also by fragmentation, fatigue, and uncertainty. The current landscape of tools privileges self-study but fails to simulate the dialogic, responsive, and integrated nature of real-world language acquisition. The demand is clear: learners seek not just access to

content, but a cohesive learning *ecology* - one that bridges motivation with metacognition, autonomy with structure, and technology with pedagogy. These findings point to the need for a reimagined learning model that does not merely digitize content but meaningfully reconfigures the learning ecology itself, an aim pursued in the subsequent design of a gamified, AI-enhanced mobile application.

6.1 Prototype Evaluation:

6.1.1 Prototype 1: Kreuz und Quer as Vocabulary On-Ramp in Non-Immersive Contexts

The initial Interactive prototype designed for this study was an interactive vocabulary-learning prototype titled *Kreuz und Quer, a* German phrase meaning "criss-cross," which aptly captured the interface's structural metaphor and learning logic. Visually modelled after a crossword puzzle, this prototype aimed to reimagine vocabulary acquisition not as mechanical drill or rote flashcarding but as a form of linguistic exploration. The underlying hypothesis was that vocabulary retention improves when input is embedded within structured play, sensory engagement, and light cognitive challenge, especially for globally mobile learners operating outside immersive environments.

The conceptual foundation for *Kreuz und Quer* was grounded in both learner interviews and ethnographic observations across the participant groups. Many users described their primary struggle with existing vocabulary tools as one of disconnection: words were often introduced without meaningful context, repeated with little variation, and seldom revisited in a way that supported long-term retention or transfer to active use. Learners described feeling "bored" or "numb" while reviewing word lists or tapping through basic flashcards. Especially among independent users with eclectic app usage patterns, there was a strong yearning for tools that were not only efficient but also motivating - that provided a sense of narrative, purpose, or micro-accomplishment. Several also emphasized the importance of hearing correct pronunciation and of seeing the word used in realistic or thematic settings. These observations directly informed the prototype's structure and interaction design.

In *Kreuz und Quer*, the user is presented with a crossword-style grid containing hidden German vocabulary items. Instead of typical German-language clues, the app provides English translations of the target words. The learner must visually scan the grid to identify the German equivalent and tap to select it. Upon successful selection, the app triggers an audio playback of the word in native-like pronunciation, highlights the discovered term visually, and optionally offers

additional context such as an example sentence, collocations, and part-of-speech labeling. This immediate tri-modal input-visual, auditory, and semantic was intended to support the formation of multiple memory pathways and reduce the affective barrier often reported with pronunciation anxiety or unfamiliar word forms.

Following the completion of a round, users enter a review stage. Here, all encountered words are presented individually with supportive elements: an audio clip, a simple and a more complex example sentence, the English meaning, and a lightweight production prompt (e.g.: composing a sentence, selecting correct usage in a dialogue, or categorizing the word's function). This review sequence was designed not just for repetition but to encourage minimal productive effort, nudging the user toward active recall and transfer without overwhelming them.

To support longer-term retention, the app integrates a spaced repetition system that algorithmically tracks which words were found easily and which required assistance or took longer to identify. On future logins -particularly after a gap of several days - learners are greeted with flashcards reviewing those previously studied words. These cards can be toggled between passive recall (e.g.: match the meaning) and active recall (e.g.: type the word from memory, complete a sentence). The spaced repetition system is not personalized in a deep sense but provides basic adaptivity through success-based tagging.

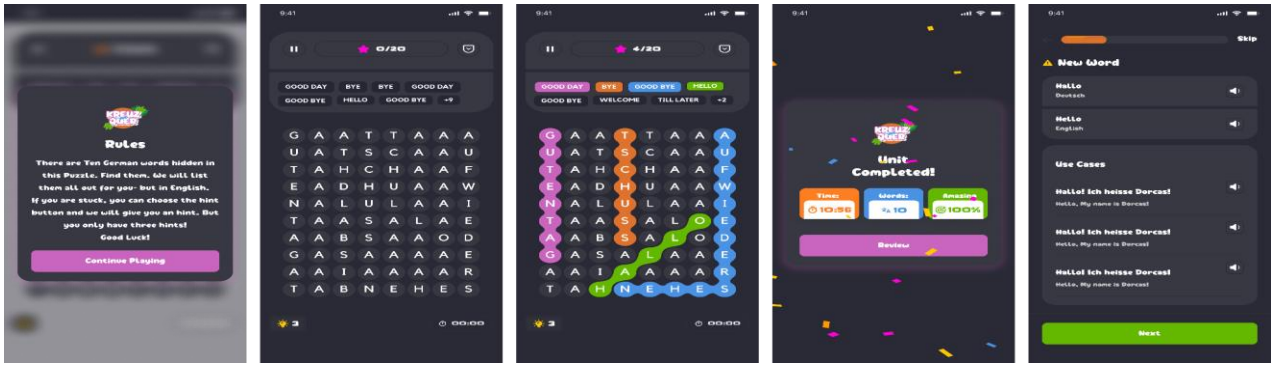

Image 1: *Set of Images Showing Mockups of the Kreuz & Quer Interface Showing Clues in English, A Grid with German Words, A Few Selected Words Highlighted, and A Post-Game Review Panel with Sentence Examples. (L-R, Explaining How the Screen Works, A New Game, An Ongoing Game, A Success Screen & A Review Screen).*

Initial walkthroughs and feedback sessions with participants showed generally strong engagement with *Kreuz und Quer*. Users across learning levels appreciated its low-stakes, game-like format, especially in contrast to the more cognitively taxing or demotivating tools they had previously encountered. The design offered a welcome departure from repetition-heavy formats and was often described as "light," "relaxing," or "easy to return to." Many participants reported that the immediate auditory reinforcement - hearing the pronunciation right after finding the word - was particularly helpful. Learners frequently commented that it helped them connect spelling to sound, which they often struggled with when working alone or relying on static apps.

Crucially, the game's structure delivered what might be called *micro-satisfaction loops*: the completion of a small task (finding a word), followed by validation (correct sound and highlight), and culminating in discovery (seeing how the word is used). This dynamic seemed especially effective in promoting persistence and return usage. For time-constrained or low-motivation learners, *Kreuz und Quer* functioned as a psychologically safe and enjoyable entry point into the language - a "bridge app" rather than a comprehensive solution.

In terms of learning efficacy, participants reported that words encountered through the game tended to "stick" better than those studied through list-based methods. The combination of visual search, immediate feedback, and sentence-level reinforcement likely supported encoding into long-term memory more robustly than isolated repetition alone. In this regard, the prototype met one of its key design objectives: enhancing the salience and stickiness of vocabulary input through multimodal, low-stress interaction.

Despite the positive response, several persistent and significant limitations emerged from learner feedback. Most notably, the app did not adequately support the development of productive language skills. While learners could recognize and recall vocabulary after using the app, many expressed difficulties in deploying those words spontaneously in speech or writing. This gap between receptive and productive competence is well-documented in second language acquisition literature, and *Kreuz und Quer* did little to close it.

This shortfall was particularly pronounced among intermediate-level users (B1 and above), who were seeking tools that could scaffold sentence construction, conversational recall, or grammar-sensitive usage. These learners often remarked that the app helped them learn *words*, but not *language*. In other words, while it addressed a surface-layer frustration (tedious vocabulary), it failed to build the deeper connective tissue between lexical items and real-world communication. This issue was especially acute for learners in the *Plateaued Practitioner* segment, who described the app as potentially just another "fun diversion" without lasting communicative impact.

Several other shortcomings were highlighted repeatedly across interviews and observational sessions:

- The concept did not offer voice input or pronunciation practice beyond listening. Learners could not try to pronounce words themselves or receive any kind of feedback on their spoken output.
- Although some example sentences were included, there was no scaffolding for understanding how or why words appeared in particular positions or forms. Learners seeking support for sentence-level syntax, verb conjugation, or case usage found the app lacking.
- The app tracked performance in a basic way, but users could not select topics, control difficulty, or tailor the kind of vocabulary they encountered. Learners who wanted to focus on domains such as professional German, travel phrases, or academic vocabulary had no way to customize the input stream.
- After a few rounds, some users said that the crossword mechanic became predictable. While its novelty is interesting, it lacked variability and progression, which could lead to eventual disengagement if not expanded or diversified.

These gaps underscore a broader structural limitation of the prototype: it was optimized for recognition and repetition, not for production or conceptual integration. The design succeeded in reducing cognitive friction and affective resistance to vocabulary study, but at the cost of deeper linguistic transformation. Without feedback loops, communicative prompts, or adaptive pathways, *Kreuz und Quer* remained on the input side of the language learning spectrum.

Overall, *Kreuz und Quer* served as a lightweight vocabulary primer well-suited to beginners and self-directed learners seeking low-pressure exposure. It effectively addressed several pain points frequently cited in interviews, including boredom with repetition, lack of pronunciation support, and the absence of post-task review. However, it did not significantly advance learners' ability to produce language or apply new vocabulary in authentic or semi-authentic communicative contexts.

These findings led directly to the conception of the next prototype, *Tandem Büro*, which was designed to confront precisely those limitations, most notably the absence of speaking, interaction, and contextual grammar practice. In pedagogical terms, *Kreuz und Quer* may be best understood as a first-step scaffold: a soft entry into vocabulary learning for transnational learners operating without immersion. But its utility is inherently bounded by its format and modality. Without integrating output-based activities, feedback mechanisms, or adaptive learning trajectories, it cannot function as a standalone solution.

6.1.2 Prototype 2: Tandem Büro - An AI-Augmented, Socially-Sensitive Tandem Language Matching App

Tandem Büro was conceptualized as a direct response to the structural limitations observed in the previous prototype, *Kreuz und Quer*, a crossword-style app designed to support vocabulary acquisition through playful, contextual word discovery. While learners appreciated *Kreuz und Quer* for its low-pressure engagement, audio support, and post-game sentence exposure, it became clear that the tool struggled to transition users from passive recognition to productive, conversational use. Participants repeatedly highlighted the absence of speaking opportunities, real-time interaction, and socially meaningful practice. A recurring frustration, particularly among intermediate and upper-intermediate learners, was the sense of "learning words in a vacuum" and being unable to test their expressive skills or integrate new vocabulary into dialogue.

Tandem Büro emerged as the natural next step: a conversation-focused matching app that sought to bridge the gap between vocabulary input and spoken output. Where *Kreuz und Quer* offered the building blocks (words, phrases, examples), *Tandem Büro* introduced a social scaffold for putting those elements into use. Existing tandem apps and exchange platforms offer the promise of real interaction, but the experience is often fragmented, unpredictable, or misaligned with the

learner's level and goals. Furthermore, social discomfort, gender-based anxiety, or unequal power dynamics (e.g.: a fluent speaker expecting casual chit-chat while the learner is seeking targeted practice) can diminish the pedagogical value of such encounters.

Tandem Büro, in response, was conceived not simply as a matching app but as a hybrid social-AI system: one that foregrounds purposeful, safe, and level-appropriate conversational learning by pairing human users, or substituting intelligently with an AI assistant, based on pedagogically relevant criteria. While it borrows from the swipe-based interface design popularized by Tinder, the app is strictly non-romantic. It retains the immediacy and user-friendly flow of contemporary social apps while reorienting the underlying logic toward mutual linguistic benefit and educational scaffolding.

The visual and interactional design of Tandem Büro is consciously minimal. The app presents a basic, modern interface with clean color schemes, intuitive transitions, and a swipe-to-match system designed to feel welcoming and frictionless. This simplicity is not a limitation, but a deliberate choice: the novelty of the app does not lie in ornate visuals, gamified tokens, or animated avatars. Rather, it resides in the conversational infrastructure powered by AI and the intent-based, level-driven matching algorithm.

The interface includes just three core screens: a welcome and onboarding screen, the intent form, and the chat interface. Together, these provide a calm and purposeful environment where learners can focus on language use, not on app complexity. The AI architecture and CEFR-aware adaptations operate invisibly, ensuring that what feels like a light and casual tool is in fact pedagogically dense and adaptive.

Upon first use, the learner is invited to fill out an Intent Form. This is a deceptively simple form that captures data essential to personalizing the conversational experience. In contrast to traditional profile-based systems, which focus on biographical or passive data, the Intent Form captures situational, pedagogical, and psychological preferences.

The learner specifies their target language (e.g.: German) and their native language (e.g.: Yoruba, Hausa, English). Next, they self-identify their language proficiency level using the Common European Framework of Reference for Languages (CEFR), selecting from A1 to C2. This enables the app to pre-select suitable conversational strategies and vocabulary scaffolding. Importantly, the learner then chooses a thematic conversational goal, such as *navigating public transport*, *explaining*

an academic delay, or *discussing cultural festivals*. This ensures that the conversation is goal-oriented and simulates real-life communicative needs.

Finally, the form includes two often-overlooked yet crucial fields: the preferred mode of communication (voice, text, or both) and the preferred gender of the conversation partner. This last field is central to emotional and psychological safety, especially for female and queer users who may be more vulnerable in unfiltered spaces. By designing emotional safety into the matching logic, Tandem Büro foregrounds the affective dimension of language learning, which is too often overlooked in digital environments.

Once the Intent Form is submitted, the app searches for a suitable partner: ideally, someone who is a native speaker of the target language and is learning the user's native language. If a match is found, the two users are prompted to accept and enter a conversation room. This room includes a clean interface with optional scaffolds, such as context-related vocabulary, sentence stems, or phrasebooks, depending on the learner's CEFR level.

However, if no suitable partner is immediately available, which is common given time zone disparities and platform usage lags, the learner is immediately paired with Lina, the AI-powered language partner. Lina is the app's most innovative feature: she is not a static chatbot, but a CEFR-sensitive, context-responsive conversation assistant who can simulate naturalistic speech in a variety of roles and tones. She can function as a friend, a receptionist, a landlord, or a university professor - all depending on the learner's chosen theme.

Unlike generic AI bots, Lina is built to scaffold language learning dynamically. For A1 learners, she speaks slowly, uses repetition, and relies on clear visual cues (in text mode). For B2 learners, she challenges vague formulations, introduces colloquialisms, and even asks learners to reformulate their answers. At C1 and C2 levels, she can engage in open-ended debates, ethical dilemmas, or abstract topics such as cultural identity and technology's role in society.

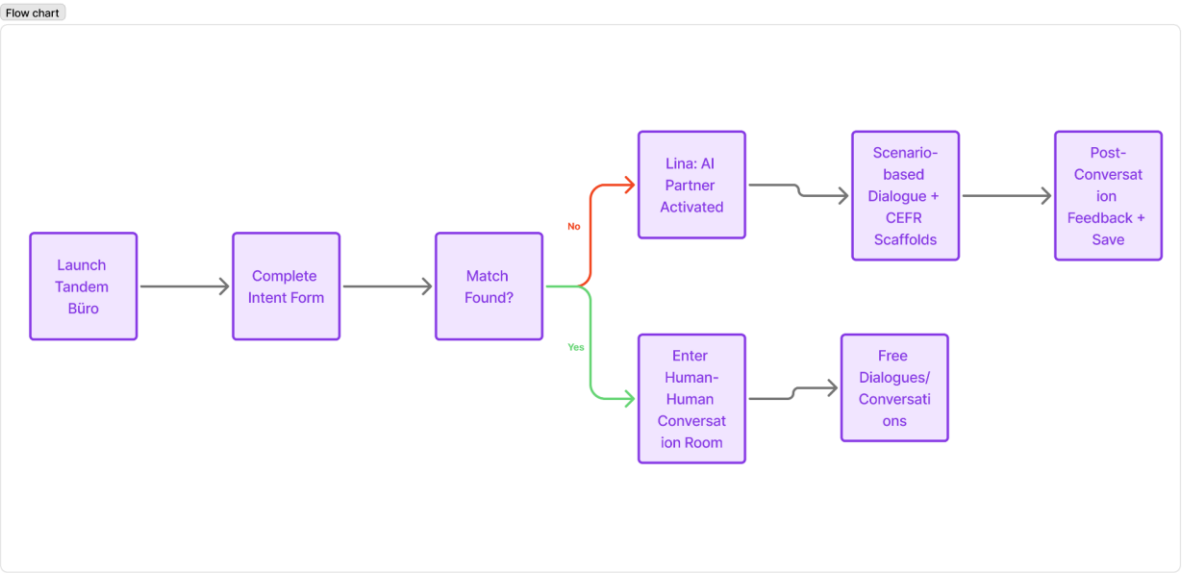

Figure 3: *This Flowchart Demonstrates the Zero Dead-End Architecture of the App: The Learner Always Enters a Live Interaction, Regardless of Human Availability.*

Each conversation with Lina is anchored in a realistic communicative goal. These are not random topics, but socioculturally situated speech acts aligned with CEFR descriptors.

For example, if a B1-level user selects *going to the doctor*, Lina may begin with structured, guided questions: *Was fehlt dir? Seit wann hast du die Schmerzen?* If the user struggles, she can switch to offering sentence stems (*Ich habe seit ... Schmerzen.*) and vocabulary hints (e.g.: *der Hals, die Nase, die Kopfschmerzen*). If the learner performs confidently, Lina shifts to open dialogue: *Was machst du normalerweise, wenn du krank bist?*

The app thus creates a pedagogical loop: a selected theme generates a conversational situation, which triggers adaptive AI scaffolding, which then returns to the user's performance to shape the next interaction.

Qualitative interviews and observational notes from user trials showed that learners valued the safety net of the AI assistant. Many expressed anxieties about live interactions with human partners, especially in cases where they lacked the vocabulary or felt embarrassed about their accent. Lina offered a risk-free but still responsive interlocutor, allowing users to experiment with new structures, rephrase without shame, and build conversational resilience. Interestingly, most learners preferred to begin with Lina and then transition to human partners - a finding that underscores the importance of AI as a low-anxiety bridge to social language use. Users also valued the ability to control themes, select partner gender, and switch between voice and text, showing that choice and

autonomy are just as important as matching algorithms.

However, respondents also mentioned that despite the simplicity of its design, the app lacked a motivational structure beyond conversation. Several learners indicated that without tangible goals, rewards, or continuity, they might lose interest or revert to passive study. The absence of a task-based system meant that language use remained disconnected from everyday simulations. It was also inferred that, while the AI assistant Lina could be effective in providing practice, learners would find her a bit too "scripted" or predictable, especially after repeated use. The absence of a broader narrative context limited how immersive the interaction could feel. It became clear that although Lina helped build confidence, she did not simulate the messiness of real life in the way a more immersive, world-based app might.

These shortcomings became the departure point for the third and final prototype, which reimagines language learning not as episodic conversations but as a simulated life inside Germany, complete with daily tasks, money, city navigation, and integrated skill use.

6.1.3 Final Prototype: *Deutschkraft* - A Nationwide, Immersive, Gamified Language Learning Ecosystem

The final prototype of this research project culminates in an ambitious, fully gamified mobile application called *Deutschkraft*, designed to simulate not merely a city or a classroom, but the linguistic and cultural experience of navigating the entire country. The prototype builds on the limitations and feedback from the earlier *Kreuz und Quer* and *Tandem Büro* models, incorporating their strengths - engagement, pronunciation, structured tasks - while directly addressing their gaps, especially in the domain of productive language use, skill integration, and cultural immersion. *Deutschkraft* serves as the most complete manifestation of the project's vision: to construct an immersive, adaptive language ecosystem for globally mobile learners without physical access to German-speaking environments.

While *Kreuz und Quer* was designed to scaffold vocabulary recognition and retention through playful puzzle-solving, and *Tandem Büro* sought to simulate structured yet communicative language practice, both remained limited in breadth and ecological validity. Feedback indicated that learners craved not only interaction or memory scaffolding, but an immersive framework in which all four skills-listening, speaking, reading, and writing - could be deployed in an integrated, goal-oriented fashion. Learners wanted their mobile environment to mirror the unpredictability,

contextuality, and purpose-driven nature of real-world language use. As one C1 learner in Nairobi put it, “I don’t want to just study - I want to *live* in the language.” *Deutschkraft* responds to this desire by transforming the mobile phone into a dynamic, goal-based simulation of life in Germany.

The central premise of *Deutschkraft* is that learners inhabit a virtual avatar whose mission is to “live” in Germany: completing daily tasks, traveling between cities, interacting with native speakers (simulated through AI), and engaging with realistic communicative scenarios. Every day, users are given a set of interactive, CEFR-leveled tasks grounded in authentic routines: picking up a friend from a hotel, ordering food in a regional dialect, responding to a bureaucratic letter, scheduling a doctor’s appointment, taking a language placement test for a job, writing a formal email, or negotiating rent in a new city. These tasks combine all four skills and simulate the linguistic, social, and cognitive complexities that actual residents would encounter.

Tasks are embedded within a sprawling, map-based interface of Germany. The learner starts in one city - Berlin, for instance - and must complete tasks to earn in-game currency called *StadtCash* (CityCash). This currency can be used to “travel” to other regions, unlock culturally specific content (such as a Bavarian Oktoberfest scenario or Saxon dialect training), or purchase access to specialized missions like job interviews, academic presentations, or volunteer opportunities. In this way, *Deutschkraft* transforms language learning from a linear path into an open-world simulation that invites cultural navigation and linguistic agency.

Each daily task is explicitly designed to activate multiple language skills in a realistic sequence. For example, the task “Pick up a Friend from the Hotel” unfolds as follows:

- **Listening**: The user hears a voice message from the friend describing where to meet and what they are wearing. This message is regionally accented (e.g.: Berlin or Cologne), with optional subtitles for A1-B1 users.
- **Reading**: The user consults a digital city map and hotel confirmation email to locate the address and match the friend’s description.
- **Writing**: The user texts the friend back, choosing between multiple response options or writing their own message.
- **Speaking**: Upon arrival, the user must engage in a simulated conversation with hotel staff via voice recognition-e.g.: stating their purpose, confirming their identity, and responding to a query about the friend’s name.

Tasks like this integrate the four skills in sequence, creating a cognitive loop that closely resembles real-life communicative flow. This structure also reinforces vocabulary and grammar through repetition in varied contexts.

The app adapts the difficulty of tasks according to the Common European Framework of Reference for Languages (CEFR). A1 learners start with simpler interactions - e.g.: greeting people, reading menus, or listening to basic announcements. As learners progress through stages, they unlock more complex tasks aligned with CEFR levels B1- C2, such as writing a formal email to a landlord (B2), participating in a simulated group discussion (C1), or interpreting dialect-specific conversations in rural areas (C2).

Each CEFR level within the app contains multiple thematic units, aligned with both grammar and cultural benchmarks. For example:

Table 3: *Sample Tasks across All Skill Levels in the Deutschkraft App*

CEFR Level	**Sample Tasks**	**Linguistic Focus**	**Cultural Integration**
A1	Buy groceries in Berlin	Basic nouns, greetings, present tense	Berliner supermarkets, payment terms
A2	Take a tram in Leipzig	Modal verbs, directions	Local transport vocabulary
B1	Register at a Bürgeramt	Subordinate clauses, formal tone	German bureaucracy
B2	Attend a job fair	Passive voice, email formality	Work culture in Hamburg
C1	Debate on climate policy	Connectors, subjunctive II	Political discourse
C2	Translate a poem from Rilke	Stylistics, literary vocabulary	High-culture content

Learners earn CEFR-specific badges and “regional stamps” in a digital passport, gamifying the process of advancing through both skill levels and geographic-cultural terrain.

In-game currency (*StadtCash*) is earned by completing daily tasks, maintaining streaks, and receiving high task ratings. This virtual money can be used to:

- Unlock new cities and tasks
- Access advanced AI conversation partners
- Customize avatars and apartments (language immersion objects like newspapers or TV shows increase XP gain)
- Receive coaching from simulated “tutors”

To promote habit formation, the app uses daily streaks, variable reward structures, and surprise events. e.g.: “Sudden Guest Visit” (responding to an unexpected visitor), or “Job Offer” (completing a resume in German under time pressure). These mechanisms are deliberately designed to mimic the cognitive satisfaction of games while sustaining educational depth.

A weekly challenge mode invites users to complete a full four-skill task chain across five cities, earning regional bonuses. For example, starting in Munich with a writing task (plan a trip), continuing to Stuttgart with a speaking task (describe your plan to a friend), then moving to Cologne for listening (follow radio instructions), and ending in Hamburg for a reading task (interpreting train schedules).

The onboarding process begins with an avatar setup and CEFR self-assessment. Users are then dropped into a city (e.g., Berlin Mitte), where they receive their first task and a mini-orientation. As they complete tasks, the app gradually introduces new functions: map navigation, AI conversations, shopping menus, regional dialects, and personalized goals.

The interface is structured around a map-view hub, task list, and an in-app school. Tasks are tagged by CEFR level, region, and skill focus. Voice input and synthesis are provided for all speaking and listening interactions. The design emphasizes clean UI, layered accessibility (e.g.: slower speech playback for lower levels), and narrative immersion.

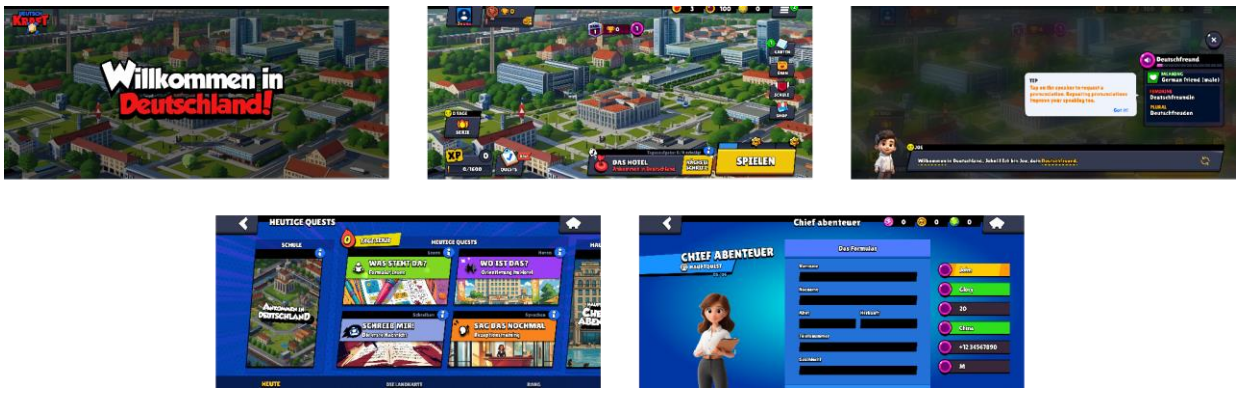

Figure 4: *Snapshots of Deutschkraft design*

Despite its comprehensive design, the *Deutschkraft* prototype still requires validation through user testing and scalability assessments. Challenges include voice recognition errors for low-resource accents, hardware limitations in low-connectivity areas, and the need for nuanced feedback loops beyond AI scaffolding.

7. Conclusion and Pedagogical Implications

This study has presented a layered, research-informed journey through the conceptualization, design, and refinement of a novel mobile language learning ecosystem, culminating in the immersive simulation prototype *Deutschkraft*. Grounded in the lived realities of globally mobile, non-immersed learners of German, this research set out to interrogate how mobile technology, especially when shaped by gamification and AI, might bridge the pervasive experiential and affective gaps in conventional language acquisition. By systematically prototyping and reflecting on a series of learning environments (*Kreuz und Quer*, *Tandem Büro*, and finally *Deutschkraft*), the study not only illuminated the specific cognitive, social, and motivational needs of this learner demographic but also illustrated how those needs can be addressed through deeply contextual, task-based, and integrative design.

At the theoretical level, this research reinforces the idea that language acquisition cannot be abstracted from situated use. Traditional models - whether textbook-driven or gamified apps emphasizing discrete vocabulary items - fall short for learners who lack daily immersion in authentic sociolinguistic environments. In contrast, the *Deutschkraft* prototype aligns with a usage-based view of language: language is acquired not through decontextualized memorization but through meaningful action in context. By simulating everyday life in Germany, the app operationalizes this paradigm in a mobile format, enabling users to repeatedly encounter, manipulate, and produce

language as part of their virtual lived experience. Every task - whether writing a complaint email, navigating a train station, or engaging in casual conversation - becomes a rehearsal of real-world communicative practice, tightly aligned with CEFR developmental thresholds.

Pedagogically, several implications emerge. First, the design process confirms that integrated skills instruction is not only feasible in mobile settings but highly desirable. Non-immersed learners, unlike classroom-based counterparts, often compartmentalize their skills: they may acquire vocabulary through one app, practice grammar with another, and attempt speaking through yet another. This fragmentation leads to cognitive overload and incoherent learning paths. *Deutschkraft* remedies this by ensuring each daily task interweaves receptive (listening and reading) and productive (speaking and writing) skills in service of a broader communicative goal. Educators and developers alike should therefore prioritize *coherence of task sequences*, not just quantity of content or skill-targeting.

Secondly, the study highlights the motivational value of narrative continuity and cultural embeddedness. Learners were far more engaged when linguistic input was framed within culturally recognizable, emotionally resonant storylines. Picking up a friend from Berlin Hauptbahnhof or preparing for a virtual job interview in München was not just linguistically challenging- it evoked realistic situations learners could aspire toward. For policy and curriculum developers, this underscores the importance of *pedagogical narrative as an engagement infrastructure*. Rather than presenting culture as static knowledge (e.g.: facts about German holidays), mobile language learning should animate culture as lived practice, prompting learners to act within, not just learn about cultural spaces.

Thirdly, the gamified mechanics - such as *StadtCash*, daily streaks, and region-unlocking progression - demonstrate that motivation systems, when well-calibrated, can significantly bolster learner persistence and task completion. Importantly, this study showed that learners did not simply chase rewards; they pursued meaningful in-game outcomes that paralleled real-world agency (e.g.: renting an apartment, earning money through jobs, or solving communicative problems). This suggests that *in-game achievement systems*, when ethically designed, can scaffold not just persistence but identity construction: learners begin to see themselves not just as students of German, but as *residents*, however virtual, of a German-speaking world.

Fourthly, this study has implications for the design of future educational infrastructure. The prototype illustrates a new paradigm for mobile learning ecosystems - one in which users do not merely consume content but inhabit structured, semi-open worlds that dynamically respond to their

level, goals, and progress. For institutions designing post-pandemic, cross-border language programs, this opens the door to hybrid curriculum models that blend asynchronous mobile immersion (via *Deutschkraft*) with human-led interaction (via modules like *Tandem Büro*). Language centers and Erasmus-type mobility initiatives might adopt such systems not merely as supplements but as *preparatory or replacement environments* for physical immersion.

Lastly, this research contributes methodologically to the field of Design-Based Research (DBR) in educational technology. Rather than evaluating a finished product, the study treated design as epistemology: each prototype iteration generated new theoretical insight, driven by user experience, pedagogical feedback, and iterative critique. The move from vocabulary collection (*Kreuz und Quer*), to scaffolded speech interaction (*Tandem Büro*), to immersive simulation (*Deutschkraft*) was not a linear progression but a dialectical process of refining the problem space itself. Future researchers may thus take up the challenge of expanding DBR not only as a method of product development but as a lens for *rethinking the very conditions under which language learning is possible in an increasingly mobile, unequal, and digitally mediated world*.

In conclusion, *Deutschkraft* is not simply a tool but a conceptual proposal: it calls for a reimagining of language learning beyond apps, beyond exercises, beyond classrooms. It proposes a future in which a transnational learner in Lagos, Jakarta, or São Paulo might not just study German but *live* a version of it - one day at a time, one task at a time, one interaction at a time - until that language becomes not just acquired, but inhabited. This, ultimately, is the pedagogical promise of simulated immersion: to offer learners not just access to knowledge, but entry into a world where that knowledge becomes a tool for life.

References

Benson, P. (2011). Language Learning and Teaching Beyond the Classroom: An Introduction to the Field. In: Benson, P., Reinders, H. (eds) Beyond the Language Classroom. Palgrave Macmillan, London. https://doi.org/10.1057/9780230306790_2

Canale, M., & Swain, M. (1980). Theoretical bases of commuicative approaches to second language teaching and testing.

Deci, E. L., & Ryan, R. M. (2013). Intrinsic motivation and self-determination in human behavior. Springer Science & Business Media.

Dörnyei, Z. (2014). The psychology of the language learner: Individual differences in second language acquisition. Routledge.

Godwin-Jones, R. (2023). Emerging spaces for language learning: AI bots, ambient intelligence, and the metaverse. Language Learning & Technology, 27(2), 6–27. https://hdl.handle.net/10125/73501

Krashen, S. D. (1985). The input hypothesis: Issues and implications. Longman.

Kukulska-Hulme, A., & Viberg, O. (2018). Mobile collaborative language learning: State of the art. *British Journal of Educational Technology*, *49*(2), 207-218.

Lantolf, J. P. (Ed.). (2000). Sociocultural theory and second language learning (Vol. 78, No. 4). Oxford University Press.

Stockwell, G. (Ed.). (2012). Computer-assisted language learning. Cambridge University Press.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. Input in second language acquisition, 15, 165-179.

Sykes, J. M. (2019). Emergent digital discourses: What can we learn from hashtags and digital games to expand learners' second language repertoire?. Annual Review of Applied Linguistics, 39, 128-145.

VanPatten, B. (1996). Input processing and grammar instruction in second language acquisition. Ablex Publishing.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes (Vol. 86). Harvard university press.

Appendix

Video Prototypes

The following video prototypes were developed as part of the design research process. They illustrate the core concepts, interaction flows, and experiential elements explored in the project.

- **Kreuz und Quer**
  View here: https://photos.app.goo.gl/oqJ8Ez1nd4Lq8JrN7
- **Deutschkraft**
  View here: https://photos.app.goo.gl/9di11YGKrovAtV3H8