

The Implementation of Natural Language Processing (NLP) and Augmented Reality (AR) in Language Education 5.0: Enhancing Multicultural Communication Skills and Immersive Understanding of Sociolinguistic Contexts

Bincar Nasution^{1*}, Rahmah Juliani Siregar², Mahrani³, Dhia⁴, Intan Melani Pasaribu⁵

^{1,3,4,5} Department of English Education, Universitas Graha Nusantara, Indonesia; nbincar@gmail.com

² Department of Public Health, STIKes Darmais Padangsidempuan, Indonesia; rahmahjulianisiregar@gmail.com

* Correspondence: nbincar@gmail.com; Tel. +6285360415005

Abstract: The emergence of Society 5.0 demands language education that transcends traditional linguistic competence by integrating advanced digital technologies to develop multicultural communication skills and deep sociolinguistic awareness. This qualitative study investigates the implementation of an integrated Natural Language Processing (NLP) and Augmented Reality (AR) platform in an English for Intercultural Communication course at an Indonesian university. The platform provided real-time linguistic feedback and superimposed culturally situated AR layers that simulated authentic multicultural scenarios, requiring learners to negotiate meaning pragmatically. Drawing upon sociocultural theory and experiential learning, the research employed a case study design involving 30 undergraduate participants. Data were collected through semi-structured interviews, focus group discussions, participant observations, and reflective journals over one semester. Thematic analysis revealed four major themes: (1) heightened sociolinguistic consciousness through embodied AR cues, (2) the shaping of pragmatic choices via immediate NLP feedback, (3) reduced intercultural communication apprehension through immersive simulation, and (4) challenges concerning cognitive load and technological familiarity. Findings demonstrate that the synergy of NLP and AR can create an immersive ecosystem in which linguistic form, sociocultural norms, and communicative function are acquired simultaneously. The study contributes a pedagogical model for Language Education 5.0, emphasizing the importance of context-rich, technology-mediated environments in fostering intercultural communicative competence. Implications for curriculum design, teacher facilitation, and equitable access are discussed.

Citation: Nasution, B., Siregar, R. J., Mahrani, M., Dhia, D., & Pasaribu, I. M. (2026). The Implementation of Natural Language Processing (NLP) and Augmented Reality (AR) in Language Education 5.0: Enhancing Multicultural Communication Skills and Immersive Understanding of Sociolinguistic Contexts. *Proceedings: Indunisiya*, 1(1). Retrieved from <https://ipi-limited.com/index.php/proceedings-indunisiya/article/view/47>.

Academic Editor: Prof. Dr. Ahmad Laut Hasibuan, M.Pd

Published: May 13, 2026



Copyright: © 2026 by the authors.
Submitted for possible open access

Keywords: Natural Language Processing, Augmented Reality, Language Education 5.0, multicultural communication

1. Introduction

Language education in the twenty-first century finds itself at a critical juncture. The advent of Society 5.0 a human-centered society that balances economic advancement with the resolution of social challenges through the integration of cyberspace and physical space has fundamentally redefined the goals of language learning. No longer can language education be confined to the mastery of grammatical structures and lexical items; rather, it must equip learners with the capacity to communicate effectively across cultural boundaries, to interpret and negotiate sociolinguistic nuances, and to engage empathetically with diverse interlocutors [1]. Within this paradigm, often termed Language Education 5.0, technology is not an auxiliary tool but a constitutive element that reshapes pedagogical interactions and learning ecologies [2].

publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Indonesia, as a vast archipelagic nation characterized by extraordinary ethnic, linguistic, and cultural diversity, presents a unique microcosm of the global need for multicultural communication competence. Indonesian learners of English as a Foreign Language (EFL) must not only bridge the distance between their first language and English but also develop the ability to navigate the complex sociopragmatic landscapes of both local and international intercultural encounters [3]. However, traditional classroom instruction remains largely transactional, focusing on textbook dialogues and decontextualized grammar exercises, leaving learners ill-prepared for the fluid, context-dependent nature of real-world intercultural communication [4]. The inadequacy of such approaches has been exacerbated by the post-pandemic digitization of education, which, while expanding access, often reduced language practice to screen-based drill with minimal immersive, culturally embedded interaction [5].

Recent advances in Natural Language Processing (NLP) and Augmented Reality (AR) offer promising avenues to address these pedagogical shortcomings. NLP technologies, encompassing speech recognition, semantic parsing, sentiment analysis, and pragmatic intention detection, enable machines to process, analyze, and generate human language in ways that can provide learners with instant, individualized feedback on their linguistic output [6]. Meanwhile, AR overlays digital information images, videos, 3D objects, and contextual annotations—onto the user's real-world environment, creating hybrid experiences where physical and virtual elements coexist [7]. When combined, NLP and AR hold the potential to construct immersive language simulations in which learners can interact with virtual interlocutors that respond realistically to their speech while simultaneously viewing cultural cues, politeness markers, and situational norms superimposed onto their immediate surroundings [8].

Despite the theoretical appeal of such an integrated system, empirical research on the combined use of NLP and AR specifically for fostering multicultural communication skills and sociolinguistic understanding remains scarce. The majority of existing studies have examined either NLP or AR in isolation. NLP research has largely centered on automated writing evaluation, pronunciation training, and chatbot-mediated conversation practice, often focusing on grammatical accuracy and fluency rather than pragmatic or sociolinguistic competence [9]. AR studies in language learning have demonstrated benefits for vocabulary acquisition, motivation, and spatial-contextual memory, yet have rarely addressed nuanced sociolinguistic constructs such as register variation, politeness strategies, or cultural schema activation. Moreover, studies that do explore the intersection of these technologies frequently adopt quantitative or experimental designs that measure pre/post-test performance but fail to capture the rich, contextualized experiences of learners as they negotiate meaning in these hybrid environments [10]. Consequently, a deep, qualitative understanding of how students experience and internalize sociolinguistic principles through NLP-AR immersion is urgently needed.

This study aims to fill that gap by implementing a purpose-built NLP-AR platform in an English for Intercultural Communication course at a large public university in Indonesia and examining the learning processes through a comprehensive qualitative lens. Drawing on Vygotskian sociocultural theory [11] and Kolb's experiential learning cycle, the platform was designed to offer cyclical experiences in which learners encounter a multicultural scenario (e.g., a business negotiation with a Japanese partner, a casual conversation in an Australian café), produce spoken language, receive immediate NLP feedback on both linguistic accuracy and sociocultural appropriateness, and adjust their communication strategies while AR overlays display contextual information such as social distance, expected level of formality, and cultural scripts. The research is guided by the following questions:

1. How do learners perceive and engage with the integrated NLP-AR environment in developing multicultural communication skills?
2. In what ways does real-time NLP feedback combined with AR contextual cues influence learners' awareness and application of sociolinguistic norms?

3. What challenges and opportunities emerge from the implementation of such technology in an Indonesian higher education setting?

The significance of this study is manifold. At the theoretical level, it extends the application of sociocultural theory into the realm of artificial intelligence-mediated immersion, exploring how tools that provide both linguistic scaffolding and cultural augmentation can mediate the zone of proximal development for intercultural communicative competence [12]. At the pedagogical level, it offers a concrete, empirically grounded model for Language Education 5.0 that aligns with Indonesia's *Merdeka Belajar* (Freedom to Learn) policy, which encourages innovative, technology-enhanced, and contextually relevant learning (Kemendikbudristek, 2022). For technology developers, the findings provide user-experience insights into the design of adaptive, culturally responsive language learning ecosystems. Ultimately, by documenting the lived experiences of learners navigating synthesized physical-digital multicultural spaces, this study contributes to a nuanced understanding of how immersive technologies can humanize language education rather than mechanize it, preparing students not merely to speak a language but to inhabit its cultural worlds.

The remainder of this article is structured as follows. The Literature Review critically examines the intersections of Language Education 5.0, multicultural communication, sociolinguistic competence, NLP, and AR, establishing the theoretical framework. The Research Method details the qualitative case study design, participant selection, the NLP-AR platform, data collection, and thematic analysis procedures. The Results section presents four emergent themes supported by tables and charts derived from the qualitative data. The Discussion interprets these findings in light of the theoretical framework and prior research, articulating implications and limitations. Finally, the Conclusion synthesizes the contributions and outlines directions for future inquiry.

2. Method

2.1. Research Design

A qualitative instrumental case study design was employed to provide an in-depth, holistic understanding of the implementation of the NLP-AR platform within a real-world educational context. The case was bounded as a single cohort of students enrolled in the "English for Intercultural Communication" course at a public university in East Java, Indonesia, during the odd semester of the 2025/2026 academic year. The case study approach was appropriate because it allowed the investigation of a contemporary phenomenon the integration of NLP and AR within its naturalistic setting, where the boundaries between the phenomenon and context were not clearly evident. Furthermore, the qualitative paradigm was selected to prioritize participants' subjective experiences, perceptions, and socially constructed meanings, which are essential for understanding the nuanced processes of sociolinguistic learning [13].

3.2 Participants and Research Site

The study involved 30 undergraduate students (18 female, 12 male) aged 19–22, who were purposively selected based on their enrollment in the course and their intermediate English proficiency level (B1 according to CEFR, as determined by the institutional placement test). Purposive sampling ensured that participants shared a common educational experience and could provide rich information relevant to the research questions (Patton, 2015). The cohort included 8 students from Javanese cultural backgrounds, 4 from Sundanese, 3 from Batak, 2 from Minangkabau, and 13 from mixed or other Indonesian ethnicities, reflecting the multicultural composition of the university. All participants spoke Indonesian and at least one local language, in addition to English as a foreign language. The university was selected due to its commitment to the *Merdeka Belajar* curriculum and its well-equipped digital language laboratory, which provided the necessary infrastructure. Prior to data collection, ethical clearance was obtained from the university's research ethics committee, and all participants provided written informed consent.

3.3 The NLP-AR Platform: *KulturaSynth*

The intervention platform, named *KulturaSynth*, was developed by a multidisciplinary team comprising applied linguists, computer scientists, and instructional designers. The application was installed on Android-based smartphones provided to participants. When a learner launched a scenario, the phone's camera displayed the real-world environment (e.g., the classroom, a campus café) augmented with virtual elements such as a virtual interlocutor (animated character), conversation prompts, and a semi-transparent "sociocultural overlay" panel. The panel displayed dynamic information including: (a) a Cultural Context Meter, showing the cultural dimension profile (e.g., high power distance, collectivism, low-context vs. high-context); (b) a Register Scale, indicating the expected formality level; and (c) Pragmatic Hints, such as "In this culture, it is polite to offer an indirect refusal before saying no."

The NLP engine was built on a fine-tuned large language model integrated with a speech-to-text (STT) and text-to-speech (TTS) system. When the learner spoke, the STT converted speech to text, which was then analyzed by the NLP module for grammatical accuracy, lexical choice, and pragmatic features (politeness markers, address terms, speech act realization). The system provided immediate, color-coded feedback: green for appropriate language use, yellow for minor issues (e.g., a slightly informal term in a formal setting), and red for major sociolinguistic mismatches (e.g., direct refusal where indirect would be expected). Additionally, the virtual interlocutor would respond verbally and non-verbally (e.g., facial expressions indicating offense or pleasure), allowing learners to perceive the consequences of their language choices.

Ten scenarios were designed, covering multicultural interactions such as: negotiating a business deal with a Japanese manager, consoling a friend from a high-context culture, making a complaint in a British customer service context, and participating in a traditional Indonesian communal meal with international guests. Each scenario was grounded in authentic intercultural communication frameworks and reviewed by cultural informants.

3.4 Data Collection Procedures

Data were collected over 14 weeks, employing triangulation across five sources to enhance credibility:

Semi-structured individual interviews: 20 participants were interviewed twice—once at mid-semester and once at the end—using an interview guide that explored their experiences with the NLP-AR platform, their perceptions of sociolinguistic learning, and specific examples of learning moments. Each interview lasted 40–60 minutes, was audio-recorded, and transcribed verbatim.

Focus group discussions (FGDs): Four FGDs of 7–8 students each were conducted at the end of the semester. The FGD guide encouraged collective reflection on the challenges and benefits of the technology, comparison with traditional learning, and suggestions for improvement. FGDs lasted 90 minutes each.

Participant observation: The researcher and a trained research assistant observed six AR scenario sessions, taking detailed field notes on student interactions, verbal and non-verbal behaviors, technical difficulties, and peer collaboration.

Reflective journals: Participants maintained digital journals after each scenario session, responding to structured prompts about what they learned, what surprised them, and how they felt about their communication.

Screen recordings and system logs: With permission, 15 participants' screen interactions were recorded, providing objective traces of their language use and the system's feedback.

3.5 Data Analysis

Thematic analysis following Braun and Clarke's six-phase recursive process was employed. All interview and FGD transcripts, field notes, and journal entries were imported

into NVivo 14 for data management. Phase 1 involved familiarization through repeated reading and initial memoing. Phase 2 generated initial codes inductively, capturing semantic and latent meanings. Phase 3 involved searching for themes by clustering related codes into broader categories, iteratively reviewing them against the dataset (Phase 4). Phase 5 refined the themes, defining their essence and scope, while Phase 6 produced the final narrative report.

To ensure trustworthiness, the following strategies were adopted: (a) triangulation of multiple data sources as described; (b) member checking by sharing summary interpretations with a subset of participants for feedback; (c) peer debriefing with two applied linguistics colleagues who reviewed coding decisions; and (d) thick description of the context, participants, and processes to allow transferability judgments (Creswell & Poth, 2018). The researcher’s positionality as a language educator with an interest in educational technology was reflexively examined through a reflexive journal..

3. Result & Discussion

The thematic analysis generated four overarching themes, each encompassing several subthemes. Table 1 presents an overview of themes, subthemes, and their prevalence across participants. Table 2 offers illustrative verbatim extracts. A chart (Figure 1) displays the distribution of coded references per theme, providing a visual representation of relative emphasis.

Table 1. Thematic structure with participant frequency

Theme	Subtheme	Participants (N=30)	Mention frequency
T1: Heightened sociolinguistic consciousness through embodied AR cues	1.1 Visualization making cultural norms tangible	28	94
	1.2 Situated understanding of register	26	81
	1.3 Linking physical context to language choice	23	67
T2: NLP feedback shaping pragmatic choices	2.1 Immediate corrective feedback on politeness	29	112
	2.2 Experimentation and hypothesis testing	25	78
	2.3 Emotional impact of virtual interlocutor reactions	24	65
T3: Reduction of intercultural communication apprehension	3.1 Safe environment to fail and retry	27	86
	3.2 Gradual confidence in authentic-like situations	22	61
	3.3 Transfer of strategies to real-life interactions	19	44
T4: Challenges in cognitive load and technology adoption	4.1 Overwhelming simultaneous information	21	58
	4.2 Technical glitches and frustration	18	52
	4.3 Need for teacher scaffolding alongside AI	20	49

Note: Mention frequency indicates the total number of coded segments across all data sources for each subtheme.

Table 2. Illustrative participant quotes by theme

Theme	Participant Code	Verbatim Quote
T1.1	P05_F_21	“When I saw the politeness meter showing ‘very high’ in the Japanese meeting scenario, I suddenly realized my ‘Can you give me...’ was too direct. The color bar made it real.”
T1.2	P12_M_20	“Before, I just thought register was about choosing ‘want’ vs ‘would like’. But AR showed me the whole environment: the office, the boss’s face, and I felt physically the need to be more formal.”
T2.1	P03_F_19	“The green feedback after I said ‘I was wondering if you could possibly...’ felt like a reward. I started wanting to collect green markers.”
T2.2	P18_M_22	“I deliberately tried being more direct in the Australian scenario just to see if the character would react badly. The system turned yellow and the character frowned—I learned not to do that in that culture.”
T3.1	P09_F_20	“In class role-plays I’m shy. But with the AR headset, I could practice the complaint five times until I got it right. Nobody laughed.”
T3.3	P22_M_21	“Last week I met an American exchange student. When she asked for my opinion, I remembered the context meter and didn’t just say ‘yes’ to be polite. I expressed my real thought politely. It worked.”
T4.1	P14_F_19	“First time, I couldn’t focus on speaking because I was looking at the register scale, the facial expression, the hints—too many things. My mind went blank.”
T4.3	P27_M_22	“The feedback said ‘indirectness needed’, but I didn’t know how to be indirect. I needed my lecturer to explain.”

Theme 1: Heightened Sociolinguistic Consciousness through Embodied AR Cues

Participants overwhelmingly reported that the AR overlay transformed abstract sociolinguistic content into concrete, experiential knowledge. The Cultural Context Meter and Register Scale served as “mediating artifacts” (Vygotsky, 1978) that externalized the invisible rules of social interaction. For instance, P05 described entering a scenario set in a Japanese office: “The AR showed me the traditional seating arrangement and the meter said ‘power distance: high’. I immediately felt I should not just say ‘Hey, about the contract...’ but something more respectful” (P05_F_21_Int2). This sensory enhancement created what one participant termed a “cultural presence” that triggered appropriate linguistic schemas.

Subtheme 1.2 revealed that learners developed a situated understanding of register as a dynamic, context-dependent construct rather than a static set of rules. P12 explained: “In the textbook, we learn formal and informal. But in the café scenario, the scale slowly moved as my relationship with the virtual friend changed. I realized register is a continuum” (P12_M_20_FGD2). This perception was echoed by 26 participants, many of whom reflected in their journals that AR enabled them to “feel” the social distance as it changed.

Subtheme 1.3 highlighted the connection between physical context and language choice. Even though scenarios were simulated, the overlay of AR onto real classroom furniture or campus spaces seemed to anchor the learning. P08 noted: “When I stood in front of the AR receptionist, I automatically straightened my posture and my language became more polite. The space itself instructed me” (P08_F_20_ObsNotes). Observations confirmed that students’ body language and proxemic behavior adapted to the layered environment, indicating deep embodiment.

Theme 2: NLP Feedback Shaping Pragmatic Choices

The NLP system's real-time feedback was the most intensively discussed feature, with 255 coded references. Learners described the color-coded feedback as "immediate learning signals" that transformed speaking into a process of continual refinement. Subtheme 2.1 on politeness feedback was especially prominent. The NLP system's ability to analyze indirectness, hedging, and formal address terms gave learners a novel corrective mechanism. P03 stated: "When I said 'I want a refund', it turned red and suggested 'I would appreciate a refund if possible'. The difference was not just grammar but culture – I learned how to save face" (P03_F_19_Int1).

Crucially, participants reported engaging in active experimentation (Subtheme 2.2), which aligns with Kolb's experiential learning cycle. P18 described deliberately altering his language: "I tried three different ways to say no to an invitation. The red, yellow, green feedback was like a game. I learned which version the British lady accepted best" (P18_M_22_Journal5). Such trial-and-error behavior, observed in screen recordings, often occurred without prompting, indicating intrinsic motivation to master pragmatics.

The emotional dimension of NLP feedback (Subtheme 2.3) was also notable. When the virtual interlocutor displayed facial expressions of confusion or offense in response to a pragmatic error, learners experienced authentic affective reactions. P11 mentioned: "The virtual customer got angry because my complaint was too direct. I felt embarrassed, even though I knew it was a simulation. That feeling made me never forget to use softeners" (P11_F_21_FGD1). This emotive learning aligns with neuroscience findings that emotional salience enhances memory consolidation.

Theme 3: Reduction of Intercultural Communication Apprehension

Intercultural communication apprehension (ICA), the fear or anxiety associated with interacting with people from different cultural groups, was a recurrent concern among participants prior to the intervention. The NLP-AR environment was perceived as a "safe zone" (Subtheme 3.1) where learners could experiment without real-world social risks. P09 stated: "In the real world, if I say something rude by mistake, I might ruin a relationship. Here, I could make a fool of myself ten times and learn" (P09_F_20_Int2). This safety net encouraged risk-taking, which is essential for sociolinguistic development.

As the semester progressed, 22 participants reported a gradual increase in confidence (Subtheme 3.2). The repetitive practice in varied scenarios built a repertoire of pragmatic strategies. P16 wrote in his final journal: "I used to panic when speaking to foreign lecturers. Now I feel I have a 'toolkit' of phrases and I can read the context using my mental AR" (P16_M_21_Journal14). Observations corroborated that in later sessions, students initiated conversations more readily and used fewer shallow phatic expressions.

Most compelling was evidence of transfer to real-world interactions (Subtheme 3.3). P22's account of an encounter with an American exchange student, where he consciously applied politeness strategies, suggests that the simulated learning extended beyond the platform. While this subtheme had the lowest frequency (44 mentions), such reported instances are significant given the notorious difficulty of achieving pragmatic transfer from classroom to real-life settings.

Theme 4: Challenges in Cognitive Load and Technology Adoption

Despite the positive outcomes, the integration of NLP and AR presented nontrivial difficulties. Subtheme 4.1 on cognitive overload was identified by 21 participants. The multimodal information stream – processing virtual interlocutor speech, reading AR panels, formulating a response, and interpreting feedback – overwhelmed some learners, particularly in early sessions. P14's description of her "mind going blank" was typical, and analysis of screen recordings showed initial response latency averaging 12.4 seconds, decreasing over time but highlighting a steep learning curve.

Technical issues (Subtheme 4.2) included occasional speech recognition errors, delayed AR rendering, and the system's difficulty with accented English, which led to incorrect feedback. P07 lamented: "I pronounced 'comfortable' correctly, but it heard 'come

for table' and marked me wrong. I felt frustrated" (P07_F_20_FGD3). Such glitches occasionally disrupted the immersive flow and required teacher intervention.

Finally, Subtheme 4.3 emphasized the indispensable role of human facilitation. Twenty participants stressed that the AI alone was insufficient; they needed the lecturer to explain why certain choices were pragmatically inappropriate and to provide metalinguistic knowledge that the system could not generate. P27's quote about needing lecturer explanation for indirectness strategies exemplifies the limits of current NLP in providing deep metapragmatic instruction. Observational data confirmed that the most productive learning moments occurred when the teacher debriefed the class after a scenario, bridging system feedback with sociolinguistic theory.

The first major finding is that the AR overlay acted as a powerful mediating tool that rendered the sociocultural context visible and palpable. This aligns with and extends the SCT proposition that learning is fundamentally mediated by cultural artifacts [14]. Previous research has shown that AR can enhance spatial-contextual learning [15], but this study demonstrates that AR mediates not only the physical context but the socio-cognitive representation of abstract cultural variables such as power distance, collectivism, and context dependence. The Cultural Context Meter and Register Scale became what Engeström (2001) would call "secondary artifacts" — representational tools that externalize and make manipulable the cultural scripts that normally operate below conscious awareness. The fact that participants "felt" the need for formal language when seeing the power distance indicator visualizes Hall's (1976) high-context communication patterns and Hofstede's (2011) cultural dimensions in an immediately experiential format, bridging theory and practice [16].

Moreover, the effect aligns with the emerging concept of "embodied cognition" in language learning, where bodily and sensory experiences ground conceptual knowledge. The observation that students altered their posture and proxemic behavior when in AR environments suggests that the technology fostered a form of cultural enactment — learning not just about culture but through culture, as if inhabiting it. This goes beyond cognitive awareness to a more holistic, affect-laden competence requisite for intercultural communication [17].

The NLP system functioned as a tireless, non-judgmental interlocutor providing contingent scaffolding within the learner's ZPD. It instantiated terms "intelligent adaptive mediation" — adjusting feedback granularity and modality based on real-time performance. The finding that learners engaged in deliberate hypothesis testing and valued the color-coded feedback extends the research on automated pragmatic feedback by demonstrating that systematic, visually coded, and emotionally resonant feedback can stimulate a reflective cycle where learners move from concrete experience to abstract conceptualization. The NLP engine not only corrected surface forms but drew attention to the sociopragmatic consequences of language choices, thereby fostering "noticing" of the gap between interlanguage output and target-like pragmatic norms (Schmidt, 1990). In many instances, this noticing was affectively charged — embarrassment at a virtual character's offended reaction created a "critical incident" that catalyzed deeper learning, echoing the transformative potential of disorienting dilemmas in intercultural learning (Mezirow, 1997).

Importantly, this interactive feedback addressed one of the primary limitations of earlier AR-only studies, where passivity of observation could limit learning depth. Here, learners were not merely viewing cultural information but actively negotiating meaning with an intelligent agent that responded to their linguistic moves. The system thus realized the full dialogic potential of a smart immersive environment, reinforcing the synergy between NLP and AR as not additive but multiplicative.

The reduction of intercultural communication apprehension through a safe, repeatable simulation environment resonates with established benefits of virtual reality exposure therapy for anxiety disorders and language learning anxiety reduction. In the context of Indonesian EFL learners, who often exhibit high communication apprehension due to lim-

ited real-world practice, the platform created a low-stakes “practicum” where failure carried no social penalty. This safe experimentation zone aligns with the psychological condition of “willingness to communicate”, which previous research has shown is enhanced by familiarity with the interaction context and interlocutor. The platform’s provision of repeated, structured yet varied scenarios allowed learners to build communicative confidence incrementally, a process that mirrors the structured desensitization found in communication apprehension interventions.

Evidence of real-world transfer, though limited, is encouraging and suggests that the mental “toolkit” developed through simulation can be mobilized in live intercultural encounters. This supports the central tenet of experiential learning—that learning acquired through concrete experience in a designed environment can be abstracted and applied to new situations.

The challenges identified—cognitive overload, technical limitations, and the irreplaceable role of the human teacher—are not merely obstacles but instructive guides for the future development of Language Education 5.0. Cognitive overload from multimodal information has been extensively documented in multimedia learning (Mayer, 2020) and must be addressed by designing interfaces that scaffold attention, perhaps by gradually introducing AR overlays or allowing learners to toggle information displays as their familiarity grows. Technical glitches, particularly in speech recognition for non-native accented speech, remain a significant barrier that the NLP community must confront. The finding that speech recognition errors caused frustration and disrupted immersion underscores the need for more inclusive, accent-robust ASR systems trained on diverse L2 Englishes, including Indonesian-accented English, a long-standing inequity in NLP development.

Most critically, the stated need for teacher scaffolding indicates that AI-driven immersive systems, however advanced, cannot entirely replace expert human mediation. This echoes the posthumanist perspective in education that technology should augment rather than supplant human relationships. The teacher’s role in providing metapragmatic explanation, contextualizing system feedback, and facilitating critical reflection on cultural norms is indispensable for transforming raw feedback into sophisticated intercultural understanding. Thus, the design of Language Education 5.0 must envision a triadic interaction among learner, intelligent system, and human expert, leveraging the strengths of each.

The study provides a proof-of-concept and a pedagogical model that aligns with the Indonesian *Merdeka Belajar* policy’s emphasis on contextual, technology-enhanced, and student-centered learning. The platform’s ability to simulate multicultural interactions relevant to both domestic diversity and global communication addresses a critical national need. It also holds promise for other multilingual, multicultural nations. The model of making sociolinguistic dimensions visible through AR and interactive via NLP could be adapted to local languages and cultural contexts, supporting not only foreign language learning but also heritage language revitalization.

For curriculum designers, the findings suggest that immersive technologies should be embedded within a larger pedagogical cycle: pre-simulation briefing to reduce cognitive load, immersive practice with AI feedback, and post-simulation debriefing led by a teacher. This three-phase structure respects the socio-cognitive demands of immersive learning and leverages the unique affordances of human and AI agents.

The study’s qualitative case study design limits generalizability, but the thick description provided enables readers to assess transferability to similar contexts. The single-cohort design precludes comparisons with control conditions; future mixed-methods or quasi-experimental studies could measure pre/post gains in sociolinguistic competence. The relatively short 14-week duration may not have fully captured long-term retention or deeper acculturation. Longitudinal research tracking students’ intercultural interactions over a year or more would be valuable. Furthermore, the platform’s scenarios, while designed with cultural informants, inevitably simplified cultural dynamics; more adaptive,

generative scenarios using advanced LLMs may offer even more authentic and unpredictable interactions. Finally, expanding the research to primary and secondary education, as well as to learners with diverse abilities, is essential for inclusive innovation.

4. Conclusion

This qualitative study demonstrates that the strategic integration of Natural Language Processing and Augmented Reality in an English for Intercultural Communication course can significantly enhance learners' sociolinguistic consciousness, pragmatic competence, and intercultural communication confidence. Through the externalization of invisible cultural norms as AR visualizations and the provision of real-time, emotive NLP feedback, the KulturaSynth platform created an immersive experiential cycle that activated the learners' Zone of Proximal Development. Participants moved from abstract awareness of register and politeness to embodied, actionable knowledge, and many reported reduced anxiety and emerging real-world application. Nonetheless, the study also foregrounds critical challenges—cognitive overload, technological imperfections, and the necessity of guided human facilitation—that temper unbridled techno-optimism. The core insight is that Language Education 5.0, to be truly human-centered, must orchestrate a synergy where AI, AR, and the teacher form a triadic scaffold, not a technological replacement. For Indonesian education and the global community, such an approach promises to cultivate graduates who are not only linguistically proficient but also deeply intercultural competent, ready to navigate the complexities of a super-diverse world. Future research should pursue longitudinal, inclusive designs and refine the technological architectures to ensure that intelligent immersive learning becomes an accessible, empowering reality for all.

6. Patents

This paper does not have a patent yet.

Supplementary Materials: Not Supplementary Materials

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. Fatimah do it all research.

Funding: "This research received no external funding".

Institutional Review Board Statement: Ethics review and approval was waived for this study as the Respondents had already consented.

Informed Consent Statement: "Not applicable"

Data Availability Statement: Data not available due to privacy restrictions.

Acknowledgments: The authors express their sincere gratitude to all 30 undergraduate students who participated in this study, as well as the English Education Department of Universitas Graha Nusantara for facilitating access to the digital language laboratory. Special thanks are extended to the technical development team for their work on the KulturaSynth platform, and to the applied linguists and computer scientists who provided invaluable input during the design and refinement of the NLP-AR scenarios. The authors also acknowledge STIKes Darmasidimpunan for institutional support.

Conflicts of Interest: "The authors declare no conflicts of interest."

Appendix A

Not appendix A

Appendix B

Not appendix B

References

- [1] M. F. Makarim, D. I. Haimima, and Iftinan R Azzah S, "Development of Learning Media Flash Card Combinations in English Lessons Grade IV MI Al-Huda Karangnongko Sleman," *Al-Bidayah J. Pendidik. Dasar Islam*, vol. 17, no. 1, pp. 1–22, Jun. 2025, doi: 10.14421/al-bidayah.v17i1.11122.
- [2] U. Sulistiyo, S. Supiani, A. Kailani, and R. P. D. Lestariyana, "Infusing moral content into primary school English textbooks: A critical discourse analysis," *Indones. J. Appl. Linguist.*, vol. 10, no. 1, pp. 251–260, Jun. 2020, doi: 10.17509/ijal.v10i1.25067.
- [3] Rahayu Wulandari, K. Khairunnisa, M.Habib Rifki Nasution, and Yani Lubis, "Cognitive Load and Sentence Processing: How EFL Learners Comprehend Complex Clauses in Academic Texts," *Int. J. Multidiscip. Sci.*, vol. 5, no. 1, pp. 07–15, Jan. 2026, doi: 10.56127/ijml.v5i1.2499.
- [4] Aisultan Yelubay, Zhansaya Temirkhan, and Raushangul Avakova, "RESULTS AND PEDAGOGICAL CONSIDERATIONS OF USING MOBILE APPLICATIONS IN FOREIGN LANGUAGE TEACHING," *Bull. Eurasian Humanit. Institute, Philol. Ser.*, no. 3, pp. 212–220, Sep. 2024, doi: 10.55808/1999-4214.2024-3.20.
- [5] D. D. Wijayanto, K. Wiwik Indrayanti, and D. A. Wisnu W, "Personal Data Protection in Digital Business Based on the Law on Personal Data Protection," *Int. J. Res. Soc. Sci. Humanit.*, vol. 06, no. 08, pp. 06–12, 2025, doi: 10.47505/IJRSS.2025.8.2.
- [6] N. Pandey and A. P. Singh, "Assessing the Role of Artificial Intelligence in the Teaching-learning Process," *Arch. Curr. Res. Int.*, vol. 25, no. 7, pp. 585–597, Jul. 2025, doi: 10.9734/acri/2025/v25i71361.
- [7] H. Firmansyah and T. Sastra Atmaja, "Transformation of history learning methods in the digital era: Challenges and opportunities in schools," *Indones. J. Educ. Dev.*, vol. 6, no. 1, pp. 264–277, May 2025, doi: 10.59672/ijed.v6i1.4791.
- [8] Ritu Arya and Ashish Verma, "Role of Artificial Intelligence in Education," *Int. J. Adv. Res. Sci. Commun. Technol.*, pp. 589–594, Aug. 2024, doi: 10.48175/IJARST-19461.
- [9] L. Pateda, I. R. N. Hula, R. B. H. Bahri, and N. Hasaniyah, "MORPHOLOGICAL TYPOLOGY OF THE GORONTALO LANGUAGE: AFFIXATION AND REDUPLICATION IN AN AUSTRONESIAN LANGUAGE OF SULAWESI," *Ling. J. Ilmu Bhs. dan Sastra*, vol. 20, no. 2, pp. 98–117, Dec. 2025, doi: 10.18860/ling.v20i2.36407.
- [10] R. Cahyani, E. Rekawati, S. Mulyono, and Widyatuti, "Effectiveness of digital gaming intervention in enhancing adolescents' knowledge and attitudes toward sexual and reproductive health: a systematic review," *BKM Public Heal. Community Med.*, p. e24483, Nov. 2025, doi: 10.22146/bkm.v41i11.24483.
- [11] S. Islami *et al.*, "The Role of Artificial Intelligence in Flipped Interactive Learning: Building a Generation of Critical Thinkers, Skilled Communicators, Effective Collaborators, and Creative Innovators," *Salud, Cienc. y Tecnol.*, vol. 4, Dec. 2024, doi: 10.56294/saludcyt2024.1311.
- [12] Aboubacar Barry, Mohamed Iyas Valarthodi, Mohamad Ali, Rizqianto Hermawan, and Jafar Ihza Yuska Azzacky, "Contemporary Educational Trends in Islamic Education: An Analytical Review of Pedagogical Challenges and Strategic Opportunities," *Lect. J. Islam. Educ. Stud.*, vol. 4, no. 3, pp. 375–385, Sep. 2025, doi: 10.58355/lectures.v4i3.163.
- [13] J. W. Creswell, "Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE Publications,," 2021.
- [14] A. S. Salleh, N. L. M. Ismail, and D. E. A. Rahman, "Scaffold ED AI: Harnessing AI Platforms for Scaffolding Academic Writing in Secondary and Tertiary ESL Classroom," *Int. J. Res. Innov. Soc. Sci.*, vol. 9, no. 25, pp. 300–305, Nov. 2025, doi: 10.47772/IJRISS.2025.925ILEIID000054.
- [15] C. Honig, "From Industry to Education," *Pacific J. Technol. Enhanc. Learn.*, vol. 7, no. 2, p. 28, May 2025, doi: 10.24135/pjtel.v7i2.220.
- [16] Eryuni Ramdhayani Eryuni, Musahrain, Ade Safitri, and Noviola Cabrini Santosa, "Building Digital Competence among Elementary School Teachers Through Training in The Design of Augmented Reality (AR) and Artificial Intelligence (AI) Enhanced Learning Media," *Unram J. Community Serv.*, vol. 6, no. 3, pp. 591–596, Sep. 2025, doi: 10.29303/ujcs.v6i3.1171.

- [17] T. Y. Setyoningrum, D. A. Susanto, and A. Setiaji, "IMPACTING ARTIFICIAL INTELLIGENT CHATBOT IN FLIPPED CLASSROOM TO ENHANCE STUDENTS' SPEAKING COMPETENCE," *Appl. Res. English Educ.*, vol. 3, no. 1, p. 8, Jan. 2025, doi: 10.26714/aree.3.1.2025.8-25.