



Enriching intercultural language learning through high immersion social virtual reality

Kristi Jauregi-Ondarra^{a,*}, Maria Christoforou^b, Juul Meijerink^a

^a Department of Languages, Literature and Communication, Utrecht University, Trans 10, 3512 JK, Utrecht, the Netherlands

^b Cyprus University of Technology, 30, Archbishop Kyprianou St, 3036, Limassol, Cyprus

ARTICLE INFO

Keywords:

Virtual reality
Virtual exchange
Intercultural communication
Blocking & threading narratives
Multimodal (inter)action analysis

ABSTRACT

Virtual reality (VR) has emerged as a promising tool in language education, offering immersive, interactive, and motivational learning experiences (Dhimolea et al., 2022; Sadler & Thrasher, 2023). High-immersion social virtual reality (HISVR) further enhances these affordances by enabling real-time, embodied collaboration through avatar-based interaction (Gruber & Kaplan-Rakowski, 2020; Lee et al., 2024). This exploratory case study investigates how learners negotiate intercultural meaning in an HISVR-mediated virtual exchange (VE), using Norris's (2004, 2019) Multimodal (Inter)action Analysis and Holliday's (2016) concepts of *blocking* and *threading* narratives to frame the analysis. Drawing on a qualitative case study of two international student groups engaged in three VE tasks in AltspaceVR, data were collected through VR session recordings, student reflection diaries, and focus group interviews. The multimodal analysis examined how students used linguistic and embodied resources—such as gesture, spatial positioning, and object manipulation—to construct intercultural understanding. Findings show that while *blocking* narratives were present—often expressed through conversational rigidity, cultural distancing, or spatial dominance—*threading* narratives were more frequent. These were characterised by inclusive language use, curiosity, collaborative movement in the environment, and coordinated multimodal actions. Students leveraged the affordances of HISVR to build rapport, clarify cultural references, and manage interactional challenges. The results demonstrate that HISVR fosters a dynamic, embodied form of intercultural communication that extends beyond the capabilities of other conventional technologies like videoconferencing. The study highlights the potential of HISVR for innovating language education by supporting task-based, interculturally oriented pedagogies that integrate multimodal and spatial dimensions of communication.

1. Introduction

Virtual reality (VR) technology has made remarkable strides, prompting scholars to explore its potential as a valuable innovative resource for language teaching and learning. Recent studies have highlighted several benefits of VR, including immersive experiences, interactivity, and increased learner motivation (Barreda-Ángeles & Hartmann, 2022; Dhimolea et al., 2022). High immersion social virtual reality (HISVR) environments, in particular, have garnered attention for their immersive nature, embodied interaction (Gruber & Kaplan-Rakowski, 2020), and capacity for international collaboration through telecollaboration or virtual exchange (VE) projects. Research indicates that such projects can affect students' willingness to communicate and engage in interactions (Canto et al., 2023), while supporting

intercultural communication (Gruber et al., 2023).

In an increasingly diverse and interconnected world, cultivating intercultural communicative competence (Byram, 2021) is essential for fostering mutual understanding and inclusivity. To explore this further, our study introduces a novel approach to investigating intercultural communication within HISVR environments, drawing on Holliday's (2016) concepts of cultural *blocking* and *threading* narratives. Unlike traditional frameworks that treat cultures as rigid, isolated “blocks,” Holliday's model emphasises the fluid interplay of shared cultural elements and the dynamic construction of cultural identities through narrative *threads*.

To analyse these interactions, we employ Multimodal Inter(action) Analysis (MIA) (Norris, 2004, 2019), a framework that enables detailed, moment-by-moment analysis of how meaning is jointly constructed

* Corresponding author. Utrecht University, Trans 10, 3512 JK, Utrecht, the Netherlands.

E-mail addresses: k.jauregi@uu.nl (K. Jauregi-Ondarra), maria.christoforou@cut.ac.cy (M. Christoforou), juul_1998@hotmail.com (J. Meijerink).

<https://doi.org/10.1016/j.cexr.2026.100158>

Received 2 October 2025; Received in revised form 24 April 2026; Accepted 29 April 2026

Available online 8 June 2026

2949-6780/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

through the orchestration of linguistic, visual, auditory, and gestural modalities in situated interaction. This is particularly relevant in HISVR, where communication unfolds through embodied action, spatial navigation, and interaction with virtual artefacts rather than through language alone. Through an exploratory case study of students navigating linguistic and cultural challenges in this immersive setting, we explore how HISVR can both reveal and mitigate barriers to intercultural communication.

This study makes both theoretical and pedagogical contributions. It addresses a critical gap: while existing research has focused on HISVR's impact on engagement and apprehension (Canto et al., 2023; Dhimolea et al., 2022), the multimodal affordances of HISVR remain underexplored (Chen & Sevilla-Pavón, 2023), with little known about how embodiment, presence, and avatar interaction shape intercultural meaning-making. Theoretically, it extends the application of Holliday's (2016) *blocking* and *threading* narratives to an emerging digital context—HISVR—by empirically operationalising these concepts through multimodal (inter) action analysis to examine how meaning is co-constructed through verbal, gestural, and spatial modalities. Pedagogically, it provides insights into the design of VR-based intercultural tasks, identifying how immersive, embodied interaction can support learners in navigating cultural boundaries. By linking nuanced discursive strategies to the specific affordances of HISVR, the research aims to inform more effective, inclusive, and theoretically-grounded approaches to developing intercultural competence in technology-enhanced language education.

2. Theoretical and conceptual framework

This framework synthesises the core concepts of the study: Intercultural communicative competence (ICC) in language education, VE and HISVR for intercultural language learning, and multimodal communication.

2.1. ICC in language education

In today's interconnected world, effective communication transcends mere linguistic proficiency (Schenker, 2012). ICC equips individuals with the knowledge, skills, and attitudes necessary to navigate interactions respectfully across cultures (Fantini, 2020). Encompassing affective, cognitive, and behavioural dimensions (Arasaratnam-Smith, 2017; Zhang & Zhou, 2023), ICC empowers individuals to become interculturally competent communicators (Byram, 2021).

Byram's influential model outlines five interconnected elements essential for intercultural engagement: knowledge (savoirs) of one's own and others' cultures; skills of interpreting and relating (savoir comprendre); skills of discovery and interaction (savoir apprendre/faire); critical cultural awareness (savoir s'engager); and attitudes (savoir être) such as openness and empathy. Complementing Byram, Deardorff (2006, 2009) emphasises ICC as a dynamic, cyclical process of ongoing reflection and adaptation.

Holliday's model adds an important critical lens to the work of Byram and Deardorff by challenging static and essentialist views of culture. While Byram and Deardorff focus on building intercultural competence through attitudes, knowledge, skills, and reflection, Holliday (2016) deepens this conversation by highlighting how cultural narratives can either block or enable genuine intercultural understanding. Holliday (2016) posits that the concept of “cultural blocks” might reinforce a problematic view of culture. This perspective, as described by Amadasi and Holliday (2017), constructs national cultures as static, insular entities that hinder intercultural exchange. By focusing on these perceived cultural barriers, students may oversimplify cultural differences, limiting their ability to forge authentic connections. This tendency to view cultures as monolithic entities, rather than complex and dynamic systems, can perpetuate stereotypes and misunderstandings. Holliday (2016) proposes an alternative to the “cultural blocks” perspective,

suggesting instead a “cultural threads” approach. This framework encourages participants to move beyond simplistic cultural comparisons and toward a more nuanced understanding of intercultural interactions. By focusing on shared cultural elements or threads, individuals can foster deeper connections and engage in more meaningful cultural exchange. While the block mentality is often easier to adopt, it might oversimplify cultural complexities and hinder the development of authentic relationships. In contrast, the threads approach requires participants to recognise the multifaceted nature of both their own and others' cultural identities, fostering a more holistic and reciprocal understanding.

2.2. Virtual exchange and HISVR for intercultural learning

VE connects learners from different cultural backgrounds through online, collaborative projects, offering a real-world context for developing ICC. A growing body of research underscores its efficacy in fostering ICC alongside language proficiency (Avgousti, 2018; Canto & Jauregi-Ondarra, 2017). VE projects develop communication skills (Barbosa & Ferreira-Lopes, 2023), empathy and tolerance (Chimeva & Trenchs-Parera, 2024; Çiftçi & Savaş, 2018; Lewis & O'Dowd, 2016), critical cultural awareness, pragmatic competence (Cunningham & Vyatkina, 2012), and awareness of one's own cultural relativity (O'Dowd & Dooley, 2020).

As technology evolves, HISVR has emerged as a medium that amplifies the VE experience through enhanced interactivity and presence. HISVR enables interaction through avatars in immersive digital spaces, ranging from desktop applications to fully immersive environments with head-mounted displays and haptic technology (Freina & Ott, 2015; Milgram & Kishino, 1994). This technology fosters a strong sense of presence (feeling situated in the virtual world) and embodiment (identifying with a virtual avatar) making intercultural interactions more engaging and contextually rich. Studies indicate its potential to increase motivation, engagement, and learning outcomes, particularly in vocabulary acquisition and contextualised language practice (Dhimolea et al., 2022; Lee et al., 2024; Lin et al., 2023).

While the integration of HISVR into VE remains underexplored, initial findings point to promising pedagogical value, albeit with varied outcomes. For instance, Gruber et al., (2023) examined HISVR's impact on presence, engagement, and communication, finding that while some students experienced strong spatial co-presence and social connection via avatars, engagement levels varied based on attitudes toward HISVR. In a related study, Canto et al. (2023) found that avatar-mediated anonymity reduced anxiety for some learners, fostering willingness to communicate, while others missed non-verbal cues, affecting their interaction comfort. From a teacher development perspective, Liaw et al. (2024) highlighted how HISVR can enhance educator agency and reflective practice in designing immersive VE tasks.

Collectively, these insights suggest that HISVR enriches VE by creating dynamic, multimodal environments where verbal, gestural, spatial, visual, and tactile modes converge, allowing learners to co-construct meaning in ways that extend beyond traditional text- or video-based exchanges. However, the variability in learner experiences underscores the need for further research into how the specific affordances of HISVR—embodiment, presence, and avatar interaction—actively shape intercultural meaning-making and competence development within digitally mediated exchanges.

2.3. Multimodal communication

Understanding how meaning is constructed in HISVR requires attention to multimodality. In technology-mediated language learning, communication is inherently multimodal (Calvo-Ferrer et al., 2016; Guichon & Cohen, 2016; Helm & Dooley, 2017; Satar, 2020), with meaning co-constructed through the simultaneous use of multiple semiotic resources: verbal language, images, gestures, and spatial

arrangements (Jewitt, 2014; Kress & Van Leeuwen, 2001). This multimodality becomes even more pronounced in VR environments due to their immersive and interactive nature. Learners communicate not only through language but also through non-verbal cues such as spatial movements and gestures, making multimodal interaction essential for understanding meaning-making in these contexts (Mills et al., 2022).

HISVR offers unique multimodal affordances. Unlike traditional digital platforms, it enables learners to navigate three-dimensional spaces, use gestures, and interact with virtual objects, adding layers of complexity to communication (Tan et al., 2016). These non-verbal forms of interaction—manipulating objects or using hand movements—are fundamental to meaning-making and provide learners with tools to communicate more holistically (Dhimolea et al., 2022; Mills et al., 2022).

While multimodal analysis has been widely applied to videoconferencing platforms in computer-assisted language learning research (Guichon & Cohen, 2016; Helm & Dooly, 2017; Satar, 2020; Wigham & Satar, 2021), the multimodal affordances of VR technologies remain significantly understudied. Existing VR-focused research has been limited in scope, with Palomeque and Pujolà's (2018) study of Second Life—a low-immersion platform—being one of the few to examine multimodal interaction in 3D virtual environments. Their work identified key challenges, including the management of multiple communicative channels and the intricate process of multimodal transcription.

In the context of HISVR, these research gaps become even more pronounced. While Karimi et al. (2023) theorised about immersive VR's potential for fostering multiliteracies, Chen and Sevilla-Pavón (2023) provided the first empirical evidence of negotiation of meaning processes in HISVR-mediated language learning. Applying Norris's (2004, 2019) Multimodal Inter(action) Analysis (MIA), their study revealed how learners strategically employed embodied semiotic resources—spoken language, gestures, and spatial positioning—to negotiate meaning within the HISVR environment. These findings underscore HISVR's unique capacity to facilitate complex multimodal interaction, highlighting its relevance as a platform for language education.

3. Case study: exploring intercultural meaning-making in HISVR

While research on the impact of multimodal HISVR experiences on ICC is limited, this study addresses this gap by investigating how specific multimodal aspects of HISVR contribute to intercultural meaning-making. Building upon previous research on VE and ICC (Avgousti, 2018; Canto & Jauregi-Ondarra, 2017; Çiftçi & Savaş, 2018), we employ Holliday's (2016) framework to analyse students' interactions, distinguishing between cultural *blocking* and *threading* narratives. To conduct a comprehensive analysis, MIA (Norris, 2004, 2019) is utilised. The study's main research question is: How do students engage in intercultural meaning-making within a HISVR-mediated VE setting?

This overarching question is further unpacked through three sub-questions:

- 1. To what extent do students' interactions reflect Holliday's concepts of cultural *blocking* and *threading* narratives?
- 2. In what ways does the HISVR environment itself facilitate the use of multimodal communication?
- 3. How do students utilise the multimodal affordances of HISVR to support their intercultural meaning-making?

Through an examination of these sub-questions, we intend to develop a nuanced understanding of student approaches to intercultural communication and their utilisation of HISVR's affordances in facilitating intercultural meaning-making.

3.1. Methodology

The current study¹ adopts a qualitative case study research approach (Duff, 2020). It entails the analysis of communication recordings conducted within HISVR environments, complemented by the examination of reflection diaries and focus group interviews (c.f. Meijerink, 2023).

3.1.1. Participants

The study involves two student groups: a male dyad (Group 1: CM1 and DM1) and a female triad (Group 2: CF1, DF1 and DF2), totalling five participants. Participants' ages ranged from 18 to 24 years. English proficiency levels varied from B1 to B2 for the Cypriot students (C) and from B2 to C1 for the Dutch students (D). The initiative was embedded within two courses: "English for Chemical Engineers" at the Cypriot university and "ICT and Language Education" at the Dutch university. Students were randomly assigned to dyads or triads and had no prior acquaintance.

3.1.2. Procedure

The VE sessions were conducted using AltspaceVR, a social VR platform (no longer available). Participants utilised Oculus Quest and Oculus Quest 2 headsets to immerse themselves in the HISVR environment. Prior to the sessions, participants received tutorials on headset operation and navigating the AltspaceVR app. Throughout the study, participants completed three tasks, one weekly, within the HISVR environment. Participants maintained personal reflection diaries, and participated in focus group interviews at the end of the project.

3.1.3. The HISVR-mediated VE tasks

Students participated in three communication tasks aimed at fostering open discussion and cultural exchange. These tasks were constructed based on principles of task-based language teaching (TBLT) (Ellis, 2003), promoting authentic language usage reflective of real-world communication. Task 1, dubbed "Getting to Know Each Other," facilitated the exchange of personal information and images to help students familiarise themselves with their partners. Task 2, titled "Exploring Partner Universities and Towns" encouraged students to navigate partner university premises and surrounding areas entering 3D video spaces, facilitating discussions on cultural issues. Task 3, devised by Dutch students, as part of the course requirements, centred around the organisation of an international party to share cultural insights on music, attire, cuisine, and beverages, ultimately leading to the co-creation of an international event. These tasks were designed to capitalise on VR's key affordances, including embodied interaction, spatial presence, and navigation for collaborative learning. Furthermore, VR's support for avatar-mediated interaction and multimodal expression aimed to enable learners to co-construct meaningful interactions within the immersive space (Jauregi-Ondarra et al., 2024).

3.1.4. The data

Data comprises recordings of the three interaction tasks (Table 1), reflection diaries, and focus group interviews conducted by authors 1

Table 1
Corpus task recordings 1, 2 and 3.

Group	Task 1	Task 2	Task 3	Total
Group 1	38:50 min.	45:51 min.	46:22 min.	131:03 min
Group 2	52:59 min.	60:19 min.	60:52 min.	174:10 min
Total	91:49 min	106:10 min	107:14 min	305:13 min

¹ This study is situated within the European funded E-LIVE project (<https://sites.google.com/view/eliveproject/home>).

and 2. HISVR tasks were recorded from a third-person avatar perspective by a student who participated solely for recording purposes, remaining silent to avoid interfering with interactions.

Participants completed reflection diaries immediately following each HISVR session in their preferred language (Dutch, English, or Greek). These structured diaries employed a mixed-format design across six thematic sections: (1) general impressions, (2) HISVR experience, (3) intercultural aspects, (4) communication dynamics, (5) partner cooperation, and (6) future applications.

At the project's conclusion, online focus group interviews were conducted via Zoom and Microsoft Teams in participants' preferred languages. Participants were randomly selected from each group for sessions lasting approximately 40 min each, audio-recorded for analysis. Interviews explored embodied learning experiences through questions such as: "What were the most positive and negative aspects of HISVR as a learning environment?", "Which HISVR tasks did you like the most or least, and why?", "How did the HISVR environment influence your communication?", and "How did you feel about meeting someone as an avatar in HISVR?"

3.1.5. Data analysis

3.1.5.1. Operationalisation of blocking and threading narratives. Holliday's (2016) concepts of *blocking* and *threading* narratives have not been systematically operationalised in empirical research. While Holliday describes how discourses obstruct or facilitate intercultural understanding, his work lacks a detailed methodological framework for identifying these narrative types in discourse data. To address this gap, the present study developed a coding scheme through thematic analysis (Braun & Clarke, 2006), using Holliday's constructs as a deductive framework. This operationalisation builds upon the first empirical application of Holliday's concepts in a digitally mediated context (Jauregi-Ondarra et al., 2024).

Blocking narratives were operationalised as discursive moves emphasising cultural boundaries and maintaining intergroup distinctions, including: (1) establishing differences; (2) maintaining differences; (3) reinforcing differences; (4) conversational rigidity; (5) claims of national or cultural superiority; (6) use of stereotypes; and (7) othering discourse through "we versus them" constructions. *Threading* narratives—those fostering intercultural engagement—were operationalised through: (1) bridge-building strategies; (2) establishing common ground; (3) conversational flexibility; and (4) intercultural orientation marked by curiosity and openness.

To enhance validity, the coding process was collaborative. One researcher analysed recordings and identified thematic categories, followed by team discussions until consensus was reached, ensuring analytical consistency.

3.1.5.2. Multimodal Inter(action) analysis. The study adopted MIA, which conceptualises communicative modes as integral to the individual (Norris, 2016). MIA focuses on how social actors use various modes to perform mediated actions—classified as *lower-level*, *higher-level*, and *frozen actions*—serving as primary units of analysis (Pirini et al., 2018).

Lower-level actions are the smallest pragmatic meaning units mediated by communicative modes (Chen & Sevilla-Pavón, 2023). In this study, they included visual actions (avatars), aural actions (e.g., DF1 playing Greek music), gestural actions (e.g., pointing at virtual objects), spatial actions (e.g., navigating between spaces), and tactile actions (e.g., grasping virtual objects). Each action was coded as a lower-level unit contributing to interaction.

Holliday's (2016) *blocking* and *threading* were treated as *higher-level actions*, arising from sequences of lower-level actions that collectively perform communicative functions. *Threading* moments emerged when students combined gestures, verbal explanations, and spatial navigation to build shared understanding. *Blocking* moments occurred when

lower-level actions emphasised cultural or communicative differences.

Frozen actions are higher-level actions embedded in objects or the environment, reflecting prior interactions that influence current behaviour (Norris, 2019). In this study, HISVR objects such as the basketball court served as frozen actions: prior engagement shaped subsequent interactions, either facilitating *threading* or reinforcing *blocking*.

Finally, *space proxemics* (Chen & Sevilla-Pavón, 2023) was applied to examine how participants used movement within HISVR to assert spatial dominance (*blocking*) or engage collaboratively (*threading*).

3.1.5.3. Transcripts. Analysis of HISVR recordings began with automated audio transcription using Otter.ai, manually verified and corrected to ensure accuracy—a critical step given challenges of constructing reliable multimodal transcripts (Bezemer, 2014). To systematically analyse instances of *blocking* and *threading* (Holliday, 2016), a detailed transcript captured linguistic exchanges alongside embodied interactions, spatial dynamics, and contextual factors. The first column marked temporal segments; the second identified participating avatars; the third and fourth columns—dedicated to context and mode—required repeated revisiting of video data to align observed intercultural meaning-making with specific modes. Close examination revealed how body positioning established power asymmetries, how gestures alternated between clarifying and confrontational functions, and how spatial dominance became a tool for intercultural meaning-making. These non-verbal modes were analysed in conjunction with linguistic data in the fifth column, enabling layered interpretation of how *blocking* and *threading* manifested across communicative dimensions.

3.1.5.4. Reflections and focus group interviews. Reflection diaries and focus group interviews triangulated findings from multimodal interaction analysis, capturing individual and collective reflections on students' embodied experiences, emotional engagement, and communicative challenges. A deductive approach guided coding and pattern identification, drawing on Holliday's (2016) *blocking* and *threading*, while remaining open to emergent sub-themes related to affect, embodiment, and perceived HISVR affordances. Two researchers independently coded the data, resolving discrepancies through iterative discussion to refine category boundaries and ensure analytic consistency, strengthening credibility and supporting triangulation of multimodal and reflective data sources.

3.1.6. Ethical considerations

The study received formal ethical clearance from the University's Ethics Review Board. All participants provided explicit, informed consent, with a detailed information sheet outlining the study's aims, procedures, and right to withdraw without consequence. All personal identifiers were removed during transcription, and participants are referred to by pseudonyms. Video recordings and anonymised transcripts were stored on a secure, password-protected server accessible only to the research team, in compliance with the Institution's GDPR guidelines. Raw data will be retained for ten years following study completion before being securely deleted.

3.2. Results

The results section is organised into three main subsections, each addressing a different dimension of the study's findings. The first subsection examines the linguistic intercultural meaning-making narratives that emerged during the HISVR interactions. The second explores how multimodal elements contributed to these intercultural exchanges. The third subsection presents the analysis of students' experiences, drawing on data from reflection diaries and focus group interviews to provide additional insight into their perceptions.

3.2.1. Linguistic intercultural meaning-making narratives

The interaction analysis conducted by the three researchers identified several linguistic instances of *blocking* and *threading* in the VE recordings. All *blocking* and *threading* categories, as defined in the method section, were observed. The analysis revealed that *threading* narratives (Table 3) consistently outnumbered *blocking* instances (Table 2) across all three tasks in the recorded interactions.

3.2.1.1. Blocking narratives. Numerous linguistic instances of *blocking* narratives emerged during the exchanges (Table 2). The most commonly identified *blocking* category was “establishing differences”. Participants frequently highlighted the cultural differences between their cultures for instance when discussing their daily and academic life (DM1: *And, yeah, well, we've, at our university, we only have two courses at a time. And that's a really big d(h)ifference. It's only, I go to the university three times a week. And it's only from on Mondays from nine to 11. So that's a big difference from nine to 6:30.*). Occasionally, this was followed by the participants maintaining the differences by only agreeing to the differences, instead of moving towards possible cultural threads (CF1: *“Yeah, that's a huge difference, actually”*). Additionally, these cultural blocks were, at times, enforced by repeatedly highlighting the established differences. Other *blocking* categories, such as national superiority claims (DF2: *“we are very modern ... it's all very well organised”*), enforcement of cultural stereotypes (DM1: *“So I was very surprised when I saw that. In the Netherlands, we always ride a bike”*), and “We vs. Them” narratives (DF1: *“I don't know what your songs are like on the radio, but ours are like very much American or English culture”*), were less frequent but still observed. Lastly, conversational rigidity, characterised by monopolising the conversation, topics and placement occurred (DM1: *“You're right. But uh, let's get back to our culture he. What do you think is a difference between/Dutch people and people from Cyprus or Greece?”*).

3.2.1.2. Threading narratives. While *blocking* did occur noticeably across the tasks, linguistic instances of *threading* narratives (Table 3) were more prevalent. Participants exhibited “conversational flexibility” by steering the conversation and reinterpreting meanings (CF1: *“Are*

Table 3

Quantity cultural threads.

Threading category	Representative quotes	Quantity			
		Task 1	Task 2	Task 3	Total
Building Bridges	<i>“Uhm I saw the buildings uhh that are your university, I guess. Yeah uhh very uh how do I say it. The buildings uh, places where I go on vacation look that wa(h)y.”</i>	62	35	66	163
Common Ground	<i>“Yeah, me too. We're similar in that aspect.”</i>	23	9	15	47
ICC Attitude	<i>“So it's very nice to be able to learn something new of a different country. And now I'm also really interested in you know, seeing Greece or seeing Cyprus, it would be really nice to experience it.”</i>	14	18	22	54
Conversational Flexibility	<i>“Because you say more freedom. It's like, you have more time left or?”</i>	25	13	21	59
Total		124	75	124	323

they, are they like sweet things? Or are they more like things for dinner? Or?”). The most commonly coded *threading* narrative was “building bridges”. Participants used bridge-building to foster understanding and connection between them. *Threading* narratives that foster open intercultural attitudes were also identified. These narratives displayed appreciation (DF2: *“That is very beautiful about your culture”*) as well as curiosity towards the other culture (CM1: *“I, I'm interested it actually (.) sounds very interesting. I would like to play that one”*), reflecting curiosity and a desire to learn more about each other's cultures).

Often, interactions began with a focus on national differences before transitioning to *threading*, where participants aligned with one another, clarified meanings, and nuanced cultural information. *Excerpt 1* offers a clear example of this shift. The interaction opens with DM1 initiating a strong *blocking* narrative by highlighting a perceived contrast between Dutch and Southern European communication styles. He suggests that Dutch people tend to be more direct - sometimes to the point of arrogance - and implies that people from Greece or Cyprus are generally less so. This kind of framing reflects what Holliday (2016) describes as a “cultural block”: a tendency to treat national cultures as fixed and homogeneous, which can lead to oversimplified views and hinder genuine intercultural engagement. CM1 responds with partial agreement (*“I can agree to some extent”*) but then redirects the discussion, proposing that “directness” is more likely a matter of individual personality than a clearly defined cultural trait. He supports this view by sharing personal experiences of encountering highly direct individuals in his own context, thereby resisting the essentialist logic of the initial framing. This move illustrates the *threading* approach described by Holliday (2016), which encourages participants to go beyond static cultural comparisons and instead explore complexity, overlap, and individuality within and across cultures. As the conversation unfolds, both participants begin to soften their positions. DM1 acknowledges that being direct *“can be good and sometimes it can be bad”*, while CM1 agrees that directness is *“a really good thing to have most of the time”* - with the important caveat, *“not always”*. This negotiated exchange demonstrates how a *blocking* narrative can give way to *threading*, allowing space for shared reflection and more relational forms of intercultural understanding. In line with Holliday's view, this co-construction of meaning exemplifies the weaving together of cultural threads, where identities are treated not as fixed or oppositional, but as dynamic and interconnected.

Table 2
Quantity cultural blocks.

Blocking category	Representative quotes	Quantity			
		Task 1	Task 2	Task 3	Total
Establishing differences	<i>“That's cold for you he?”</i>	18	19	20	57
Maintaining differences	<i>“ Oh, yeah, that's actually another difference. Like people in Greece and in Cyprus.”</i>	5	8	3	16
Enforcing differences	<i>“[A it's very, very, very,] very different. Yeah.”</i>	3	3	0	6
Conversational Rigidity	<i>“And we also needed to discuss actually we had to do that before we he told each other about our countries. But what's what's your idea of our culture?”</i>	9	5	9	23
National/Cultural Superiority or Competence claim	<i>“I think your country is meer more like, together and uhm well and also about so-solidarity.”</i>	6	1	1	8
National/Cultural Stereotype	<i>“In the Netherlands, we always ride a bike”</i>	6	2	7	15
Othering/We vs. Them	<i>“We, we Dutch people don't really do that.”</i>	4	1	3	8
Total		51	39	43	133

Excerpt 1. Group 1, task 1, showcasing a combination of narrative strategies

Speaker and Timestamp	Transcript
DM1 00:27:26	(...) I also think that Dutch people are more, because we also can talk about things that are maybe taboo for you. You can also be more direct and what we say so we don't really, sometimes we're a bit like, arrogant. You just say what you think. And I don't think people from Greece or Cyprus do that, do you think that too?
CM1 00:28:15	Ah, yes, I can agree to some extent. But (.) my take is that it's always depends on the person you are talking to like, I've met some very direct people here as well. Like they would they would talk like, anything that came to their mind, they would just like it out. That would even cause some problems at some points like if you are in a gathering, with many people will not many but like around ten. That could cause some problems. Yeah, sometimes it can also be good, I think, because people know what you think. And so you don't have to be fake or something like that. Because some people always try to please other people, we, we Dutch people don't really do that. Sometimes it can be good and sometime it can be bad. I am thinking about other differences.
DM1 00:29:05	It is a really good thing to have most of the time. Uhh the being direct.
CM1 00:29:29	Not always.
DM1 00:29:41	Yeah, not always obviously.
CM1 00:29:43	

3.2.2. Multimodal affordances for intercultural meaning-making

This subsection explores how multimodal resources shaped the construction of intercultural meaning within the HISVR setting. The recordings displayed a wide array of multimodal categories, encompassing visual, aural, gestural, spatial, and tactile modes. *Frozen actions* embedded in the virtual world, such as flying basketballs, rockets, and a polar bear, appeared alongside interactive features and influenced current actions. For example, avatars and the AltspaceVR manual or chat sometimes appeared on screen, especially with group 1. The *spatial mode* was evident when avatars moved within the environment and during interactions like playing basketball or choosing a conversation location. The *aural mode* included background noises from real-world surroundings and virtual sounds like rocket explosions. It occasionally emerged due to real-world events, such as DF1 playing a song to check familiarity with a Greek singer. Participants frequently employed *gestural modes*, including headset-mediated nodding and controller-based hand movements, to facilitate interaction. Participants touched their faces or VR glasses, gazed at other avatars, and made gestures like throwing basketballs. Hand gestures were used for detailed explanations, pointing when discussing participants, and gazing for uncertainty or seeking support. Finally, the *tactile mode*, involving physical contact with objects, was evident particularly in group 1's basketball game, where both partners tried to grasp the ball while holding the joysticks (Jauregi-Ondarra et al., 2024).

While these multimodal resources (verbal, gestural, aural, spatial and tactile) shaped the overall interaction dynamics, our analysis specifically interrogates how such affordances function within both *blocking* and *threading* narratives (Holliday, 2016).

3.2.2.1. Multimodal blocking narratives. *Blocking* narratives prominently emerged in Group 2 (DF1, DF2, CF1), with DF1 playing a pivotal role in shaping the group's dynamics from the outset. Her assertive behaviour, marked by both linguistic and spatial dominance, established a pattern that disrupted collaborative engagement constraining her partners' opportunities for collaboration. The integration of multimodal cues - such as gesture, gaze, movement, and spatial positioning - reveals how DF1's recurring actions reinforced a hierarchical structure within the group, ultimately affecting the flow of interaction.

As shown in Excerpt 2 (Fig. 1), at the beginning of Task 2 (*Exploring Partner Universities and Towns*), DF1 exhibits spatial and linguistic

dominance by stating, "*Shall we go outside because this is a bit lame place to (.)*" while looking away (head movement) and leading the group outside the apartment. The *space proxemics* of the group can be seen in Excerpt 2 which places DF1 in a position of control, potentially limiting her partners' autonomy and preventing them from contributing to decisions about the environment. DF1's physical *blocking* is further reinforced through the verbal mode, exemplified by the use of the negative connotations "*lame*," which she employs without providing any clarification or justification.

DF1's *body movements* also likely reinforce hierarchical dynamics. In Excerpt 4 (Fig. 3), she steps back and asks, "*Did you even see this place yet?*" (referring to another room in the immersive environment) while using hand gestures to point at the room. She then rushes ahead and enters first (Excerpt 5, Fig. 4), establishing a pattern that results in unbalanced participation in the interaction among the partners. This results in the partial marginalisation of her partners, who do not contribute to decisions regarding the location or environment.

In addition to exercising *spatial control*, DF1 also asserts verbal dominance, which constructs higher-level *blocking* actions. For example, in Excerpt 6 (Fig. 5), DF1 asks, "*Did you guys read the task?*" but does not genuinely seek input from her partners. Instead, she promptly asserts her own experience, stating, "*I did. I don't know where to begin.*" The verbal mode, realised through the use of exclusive language ("*you guys*" and "*I did*"), and the spatial arrangement (standing at a distance from the sitting area where the other two girls are) can potentially initiate higher-level actions that revolve around the building of barriers instead of bridges with her partners.

Task 3 (*Organisation of an international party*) begins only with DF1 and CF1 as the second Dutch student joined the conversation later. Excerpts 1-7 (for Excerpts 3 and 7, see appendices A and B) could potentially exhibit *modal intensity* (when a mode is more dominant than the others) since DF1's construct of *proxemics* mode (spatial dominance) elucidates *blocking* not only in one task, but in all three, and is a result of DF1's actions in the immersive environment (Norris, 2004).

Blocking instances continue to appear in the conversation that revolves around the topic of traditional snacks Dutch and Cypriot people eat at parties or family gatherings (Task 3). Excerpt 8 (Fig. 7) exhibits *modal complexity* which refers to how different communicative modes are co-dependent and coexist to construct and make higher-level actions possible (Norris, 2004). In this case, modal complexity is realised through CF1's verbal and visual portrayal as a way of facilitating understanding between her and her Dutch partner. Excerpt 8 (Fig. 7) shows CF1 trying to describe *loukoumades* in detail ("*We have something called loukoumades, which is fluffy small balls, like bread, but they are fried, and they are very sweet*"), using hand gestures to illustrate the frying process and even uploading an image of *loukoumades* via her HISVR controllers (object handling) for her Dutch partner to have a visual representation of the snack (Excerpt 9, Fig. 8).

The verbal, gestural and visual modes are intertwined in this multimodal sharing to facilitate a *threading* interaction as CF1 endeavours to elucidate as much information as she can on this cultural element and ensure her partner comprehensively understands it. However, we can observe a lack of multimodal interplay in DF1's elucidation of *bitterballen* (a Dutch, meat-based snack) compared to CF1's portrayal of a traditional snack. DF1 expects that CF1 should be familiar with *bitterballen* and does not try to describe how it looks like or how it is made, which is followed by her subsequent disappointment when CF1 does not know it ("*I haven't heard of any actually*"), as illustrated in Fig. 9 (see Appendix B, Excerpt 10). This cultural difference is accentuated by DF1's final remark ("*I don't know any snacks from your country either*"), which results in an awkward laughter (sound) from CF1.

The analysis of these Excerpts reveals how DF1's consistent use of spatial, verbal, and gestural dominance in Group 2 established hierarchical dynamics that constrained collaboration, particularly marginalising her Cypriot partner (CF1). Through multimodal *blocking* actions - such as proxemic control, exclusive language, and unilateral decision-

making - DF1 reinforced barriers to intercultural exchange.

3.2.2.2. Multimodal threading narratives. Having established these *blocking* patterns, we now examine how *threading* narratives emerge through more balanced multimodal coordination, where participants employ verbal, gestural, and visual resources collaboratively to bridge gaps. These instances are particularly evident in Group 1's interaction, where a different set of communicative patterns emerged. Unlike the disruptions observed earlier, Group 1 consistently employed *lower-level actions* that fostered inclusion and mutual understanding. The following excerpts illustrate how multimodal resources were used collaboratively to manage and overcome potential challenges, thereby constructing *threading* narratives that supported intercultural engagement.

For example, Excerpt 11 (Fig. 10) depicts the interaction between the two students following the completion of the third task. The Dutch student inquires whether his partner is interested in visiting the Netherlands, asking, "*When you think about all the stuff we talked about in these three meetings, would you like to visit the Netherlands?*" Although the gesture alone might impede the Cypriot student's understanding if the verbal context were unclear, in this instance, it highlights the word "*three*" to underscore their three sessions, thereby enhancing the verbal message with the gestural input. Furthermore, the use of inclusive language ("*we*") signals mutual collaboration in the *threading* narrative (see Appendix C) for another example of lower-level action that promotes mutual understanding).

Excerpt 13 (Fig. 12) describes the Dutch student's struggle with explaining the process of making his favourite dish, *kapsalon*. Even though the flow of conversation is interrupted with a pause ("*you get a kind of (.)*") and uncertainty ("*how do you call it in English?*"), the linguistic ambiguity is overcome through an interplay of modes to promote mutual understanding. Specifically, DM1 makes use of hand gestures to exemplify how Dutch people wrap this food with aluminium foil and exhibits proxemics behaviour by standing very close to his partner thus being able to "look into his eyes". Collectively, these coordinated lower-level actions constitute a *threading* instance, demonstrating how multimodal alignment can facilitate intercultural connection.

Finally, Excerpt 14 (Fig. 13) demonstrates how the cultural process of making *kapsalon* ("*You get the meat, and French fries, salad, and cheese on top of it, but it's melted all together, and then you add garlic sauce*") becomes clearer when the Dutch student uses hand gestures to illustrate the various layers of food. This use of hand gestures can be seen as enriching the description and may have contributed to greater clarity for the Cypriot student, potentially supporting a more natural flow to the conversation.

3.2.2.3. Synthesis of findings. Viewed across the dataset, the sequential instances discussed above reveal recurring patterns in how multimodal affordances shape intercultural meaning-making in HISVR. Because actions unfold temporally and meaning is produced moment by moment, lower-level actions (e.g., gaze, gesture, spatial positioning, verbal choices, and interaction with virtual artefacts) were examined cumulatively to trace the emergence of higher-level actions. Across tasks, affordances related to spatial control and proxemics consistently contributed to *blocking* narratives when enacted unilaterally, reinforcing hierarchical relations and limiting participation. In contrast, *threading* narratives emerged through balanced multimodal coordination, where participants aligned verbal explanations with gesture, gaze, visual support, and shared spatial orientation. This synthesis highlights how affordances function relationally and over time, rather than as isolated resources, in the co-construction of intercultural meaning. In sum, intercultural meaning-making in HISVR environments is shaped by the dynamic interplay of verbal, gestural, spatial, visual, and tactile modes. While dominant multimodal configurations can constrain collaboration, coordinated multimodal action supports mutual understanding and intercultural exchange.

3.2.3. Participants' perceptions

Participants in both the reflection diaries and focus group interviews offered rich insights that illuminate how multimodal affordances helped shape intercultural meaning-making and communicative dynamics within the HISVR environment. Overall, the immersive setting was described as conducive to interaction and engagement. DM1, for example, shared that "*It provided a pleasant atmosphere; we initiated conversation through a basketball game,*" highlighting how shared spatial experiences in the HISVR environment facilitated relaxed social entry points - often linked to *threading* narratives. Similarly, DF2 emphasised how the interactive features of the environment supported attentiveness and camaraderie, reinforcing the potential for multimodal resources to foster fluid communication.

However, these same affordances could also contribute to cognitive overload and distraction, particularly when coordination of multiple modes was required. DF1 expressed that the immersive and sensory-rich environment occasionally felt overwhelming - an insight that aligns with her observed reliance on dominant spatial and verbal strategies in the earlier *blocking* narratives. Her subjective experience thus reinforces how an inability or unwillingness to balance multimodal engagement could lead to exclusionary communicative patterns.

Spatial orientation in the HISVR environment also emerged as a key enabler of immersive, natural-feeling communication. Participants consistently likened the experience to face-to-face interaction, with one noting, "*It felt like conversing in person, as if we were in the same room despite being in different countries*" (*co-presence*). Another remarked after watching a 3D video of the partner's University and city, "*It felt different from watching a regular video; it was as if I was physically present*" (*presence*). These reflections suggest that *proxemics* and *embodied presence* - core elements of multimodal communication - enabled participants to navigate interactions more intuitively, often supporting lower-level actions that underpinned *threading* narratives.

Participants further noted how the HISVR environment made *gestural* and *spatial modes* more visible and impactful. While the avatars lacked full facial expressivity, the visibility of body gestures, such as nodding and hand movements, was repeatedly mentioned as essential to mutual understanding. One participant explained, "*I appreciated being able to observe my partner's gestures,*" while others recognised that gestures played a key role in supplementing verbal communication. This interplay between *visual* and *gestural modes* helped establish shared meaning and supported collaborative exchanges - key to *threading* narratives. At the same time, the absence of facial cues was sometimes cited as a limitation for interpreting nuanced emotional intent, showing that while HISVR afforded many modes of communication, it also had constraints.

Finally, intercultural meaning-making was identified as a salient theme. The participants described the HISVR environment as a bridge across geographical and cultural distances, providing an embodied space for cultural sharing. Many expressed curiosity about their partner's background and reported a heightened awareness of their own communicative behaviours. For some, this prompted a reflective shift in how they navigated cultural differences and adapted their communication styles, revealing how multimodal engagement in HISVR not only enabled intercultural exchange, learning and self-awareness.

3.3. Discussion

This study investigated how students engage in intercultural meaning-making within a HISVR-mediated VE setting. The analysis was guided by three sub-questions: (1) to what extent students' interactions reflected Holliday's (2016) concepts of *blocking* and *threading* narratives; (2) how the HISVR environment facilitated multimodal communication; and (3) how students utilised these multimodal affordances to support intercultural meaning-making through *blocking* and *threading*.

In response to the first sub-question, both *blocking* and *threading* narratives emerged in students' linguistic interactions, though *threading* was more frequent. *Blocking* narratives were marked by conversational

rigidity, maintenance of national and cultural differences, and instances of othering. Students highlighted perceived contrasts in university structures or social habits, occasionally reinforcing stereotypes. One student remarked on the informality of Dutch academic life compared to their own, concluding, "*That's a big difference.*" Superiority claims—such as describing one's national systems as more "modern"—reflected hierarchical thinking. These moves correspond to Holliday's (2016) *cultural blocks*, where discourse reifies national cultures as static and distinct.

In contrast, *threading* narratives were characterised by bridge-building, clarification-seeking, and inclusive framing. Learners asked open-ended questions, expressed curiosity ("*That sounds interesting*"), and reformulated content to maintain dialogue. Particularly in Task 3, students co-constructed meaning through collaborative lexical negotiation, using phrases like "*So is it more like sweet or savoury?*" or "*We also have something like that.*" These moves created space for dynamic exchange and mutual understanding—aligning with Holliday's (2016) cultural *threading*, where learners work across perceived boundaries. *Threading* narratives increased as interactions evolved, suggesting a trajectory beyond surface-level comparisons. This supports Dervin's (2016) critique of fixed cultural categories and highlights the need to approach culture as fluid and relational.

The second sub-question explored how the immersive HISVR environment facilitated multimodal communication. The platform offered students access to communicative resources beyond speech—spatial positioning, avatar movement, gestural expression, object interaction, and environmental cues. Participants described the setting as immersive and supportive of spontaneous interaction, often likening it to face-to-face communication. Many reported that the sense of shared space made it easier to initiate conversations and build rapport. For instance, the ability to "walk around" a virtual lounge provided not only functional navigation but also social cues that influenced interaction dynamics. These affordances created fertile ground for *embodied interaction*, distinguishing HISVR from text- or video-based platforms. As Freina and Ott (2015) and Gruber et al. (2023) argue, HISVR enables a sense of co-presence that enhances both linguistic production and social engagement. Students frequently used environment-related references ("*Let's go outside*") as springboards for interaction—showing that the affordances of the space became part of the communicative repertoire. At the same time, some students experienced cognitive strain, citing the absence of facial expressions and the challenge of processing multiple sensory channels. These reflections confirm the dual nature of HISVR environments: while they provide expansive communicative possibilities, they also require careful pedagogical scaffolding (Canto et al., 2023; Gruber et al., 2023).

The third sub-question focused on how students used the multimodal capabilities of HISVR—gesture, space, visual and tactile modes—to realise *blocking* or *threading* narratives. These lower-level actions (Norris, 2004) were central to the enactment of intercultural dynamics. In *threading* interactions, learners used coordinated multimodal strategies to clarify meaning and promote engagement. For instance, CF1's explanation of *loukoumades* combined verbal elaboration with gestural demonstration and image-sharing via HISVR controls. Similarly, in Group 1, gestures illustrated layered food (e.g., *kapsalon*), while proximity and inclusive body orientation supported sustained engagement. These cases demonstrate *modal complexity* (Norris, 2019), where multiple semiotic modes work in tandem to co-construct meaning.

Spatial interactions were particularly telling. In Group 1, joint navigation—choosing locations together or sharing spaces during games—fostered mutual respect and *threading*. In contrast, Group 2 illustrated how *modal intensity* in the spatial mode could contribute to *blocking*. DF1 frequently led the group without soliciting input, using proximity and movement to direct others. While DF1's actions were predominantly associated with *blocking*, her confident navigation may also be interpreted supportively: by leading the way and introducing partners to different settings, she arguably aimed to facilitate their shared experience. These interpretations underscore the complexity of

spatial behaviour in HISVR-mediated communication, where the same actions may serve inclusive or exclusive functions depending on context (cf. Chen & Sevilla-Pavón, 2023).

The use of tactile modes, though less frequent, also contributed to intercultural interaction. In the basketball scenario, students' physical coordination in handling the ball served as a low-stakes entry point for informal dialogue and rapport-building—particularly useful in early tasks where trust was still developing. This was especially evident in Group 1, who consistently began sessions playing basketball to build rapport (Jauregi-Ondarra et al., 2024). These observations align with Chen and Sevilla-Pavón's (2023) argument that embodied semiotic resources in HISVR are central to meaning negotiation and perspective-taking. They also reinforce that negotiation of meaning in immersive environments depends on learners' ability to orchestrate multiple modes of expression in socially sensitive ways.

Due to the rich multimodal affordances of the immersive environment, HISVR-mediated interactions offered a communicative experience more dynamic and embodied than typically observed in videoconferencing settings. While previous studies (e.g., Wigham and Satar, 2021) have demonstrated the complexity of multimodal interaction in video-mediated communication, the spatial and embodied affordances of HISVR further expand these possibilities by allowing participants to navigate the environment, manipulate objects, and use physical positioning as a communicative resource. This added layer transforms the nature of engagement, making communication more fluid and co-constructed. These findings underscore the importance of attending to the full spectrum of embodied interaction when examining communication in immersive environments.

3.3.1. Methodological considerations and further research

As a small-scale exploratory case study, this research offers in-depth insights but is limited in generalisability. The project was constrained by a short duration (one month) and a small participant sample (five students). While rich in qualitative depth, the findings may not capture the full range of interactional diversity observable in longer or more culturally varied exchanges. Group dynamics—such as the difference between triads and dyads—may also have influenced participation and positioning. The potential impact of gender composition on interaction patterns was not examined. Future studies should explore these variables in more depth, over longer periods, and across varied HISVR platforms.

Rather than a limitation, the scope of the multimodal analysis should be seen as a focused analytical choice. MIA was conducted exclusively on recorded interactions within the HISVR environment—focusing on what participants saw, heard, and did as avatars. Unlike studies such as Chen and Sevilla-Pavón (2023), we did not analyse participants' real-world bodily movements outside the headset. This was a deliberate choice to prioritise modes that were mutually visible and tangible to the interlocutors during the immersive exchange. However, future research might capture both in-world and real-world embodied actions.

Additionally, the early-stage operationalisation of *blocking* and *threading* as empirical categories requires further refinement, as this study represents the initial attempt at conceptualising these theories. While the coding framework was effective for identifying key patterns, the overlapping and fluid nature of real-time communication meant that some instances defied easy classification. Future research should further develop and test these categories to enhance analytical precision.

3.4. Conclusions and pedagogical implications

This study set out to explore how students engage in intercultural meaning-making within a HISVR-mediated VE, using Holliday's (2016) concepts of *blocking* and *threading* and Norris's (2004, 2019) MIA as guiding frameworks. The findings point to the central role of both linguistic and embodied modes in shaping intercultural interaction. While *blocking* narratives—marked by conversational rigidity, assertions of

cultural difference from essentialist perspectives, or hierarchical use of space—were present, the overall interactional patterns leaned more heavily toward *threading* narratives. These were realised through bridge-building, alignment, curiosity, and mutual collaboration, both linguistically and through students' strategic use of HISVR's embodied affordances.

A key contribution lies in demonstrating how HISVR enriches the communicative experience beyond what has been possible in previous VE projects that relied on low-immersion platforms such as videoconferencing or text-based chat. Studies by Avgousti (2018), O'Dowd and Dooly (2020), and others have shown that ICC can be developed through VE, even when limited to verbal and facial cues. However, HISVR offers a more expansive communicative palette—gesture, proximity, spatial navigation and proxemics, object interaction, and ambient environmental cues—allowing participants to co-construct meaning in ways that are interactive, layered, and deeply situated. This transformation from two-dimensional screen interaction to embodied co-presence redefines what intercultural communication can look and feel like in pedagogical virtual settings. Importantly, learners used these affordances both inclusively and exclusively. *Threading* narratives were often supported by coordinated use of gesture and spatial alignment, while *blocking* narratives frequently involved dominant spatial behaviour or one-sided movement through the environment.

Our findings support and extend the work of Chen and Sevilla-Pavón (2023), who highlighted the critical role of embodied semiotics in immersive language learning. At the same time, they reinforce that immersive technologies do not automatically lead to better communication; rather, they must be accompanied by pedagogical frameworks that help students become aware of how they are using these modes and what kinds of intercultural relationships they are constructing.

From a pedagogical standpoint, integrating HISVR into language education calls for a fundamental rethinking of what intercultural communicative competence means in the 21st century. This requires acknowledging the complexity of the construct and moving beyond reductive models of cultural knowledge. As Dervin (2016, p. 72) argues, we must “move away from a perspective that lays an emphasis on facts about a ‘target culture’, presents ‘grammars of culture’ (do's and don'ts), concentrates on cultural difference, and confronts cultures and civilizations,” and instead promote more fluid, dynamic, and relational *threading* narratives. In this context, teachers must be equipped not only with technical skills to manage immersive environments but also with pedagogical expertise to guide learners through embodied, multimodal, and interculturally sensitive interactions. Educators thus take on expanded roles—as language facilitators, spatial designers, intercultural mediators, and curators of embodied learning experiences.

Moreover, the immersive nature of HISVR offers a pathway to overcome some constraints of traditional classroom-based learning. It allows learners to engage in situated (Lave & Wenger, 1991) and experiential practice (Kolb, 2014) with real communicative stakes and social complexity—bringing them closer to authentic intercultural encounters than any textbook could offer. Yet, this innovation must be implemented thoughtfully. Digital inequality, unfamiliarity with avatars, and the cognitive load of managing multiple modes simultaneously are real challenges that must be addressed through inclusive design and adequate scaffolding.

Broader implementation faces significant barriers, including the financial cost of VR hardware and the technical infrastructure required. Furthermore, effective use demands considerable teacher training, both in technical operation and pedagogical strategies for facilitating VR-mediated dialogue. A practical illustration of the volatile nature of this technology is the discontinuation of AltspaceVR, which underscores the need for sustainable, adaptable solutions or a focus on platform-agnostic instructional design principles.

In sum, this study reinforces the potential of HISVR as a valuable tool for enriching language education (Lee et al., 2024) and intercultural learning (Hwang & Lee, 2024). When paired with principled task design, critical reflection, and an awareness of embodied interaction, immersive technologies offer students not just new ways of communicating, but new ways of being and relating in intercultural contexts. As the boundaries between physical and virtual spaces continue to converge, so too do the possibilities for cultivating meaningful and embodied forms of intercultural language learning.

CRedit authorship contribution statement

Kristi Jauregi-Ondarra: Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Maria Christoforou:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Juul Meijerink:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Data curation, Conceptualization.

Statements on open data and ethics

The participants were protected by hiding their personal information in this study. They were voluntary and they knew that they could withdraw from the experiment at any time. The data can be provided upon requests by sending e-mails to the corresponding author.

Ethical approval

Informed consent was obtained from each participant within this study.

Ethics

- This study has been reviewed and approved by the Institutional Review Board at Utrecht University. All procedures were performed in compliance with relevant laws and institutional guidelines, and the appropriate institutional committee(s) have approved them.
- Informed consent was obtained from all participants within this study.
- The privacy rights of human subjects within this study are observed through secure storage of all data associated with this study in an appropriate protected repository.

Declaration of generative AI in scientific writing

During the preparation of this work the author(s) used ChatGPT in order to improve accuracy and coherence of the text. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Funding sources

The project was undertaken within the context of Erasmus + project E-LIVE: Engaging Languages in Intercultural Virtual Exchange (2022–2024).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

Fig. 2 illustrates her Dutch partner immediately following her (body movement), while the Cypriot partner trails behind (body movement). It reinforces the following/trailing proxemics of Fig. 1.

Time	Participant(s)	Context	Mode	Language
00:06	DF1	DF1 leads the way; DF2 follows next, leaving CF1 behind them.	looks away (head movement) and goes out of the room (body movement) 	<i>"Shall we go outside because this is a bit lame place to (.)"</i>

Fig. 1. Multimodal transcript of group 2 (Excerpt 2).

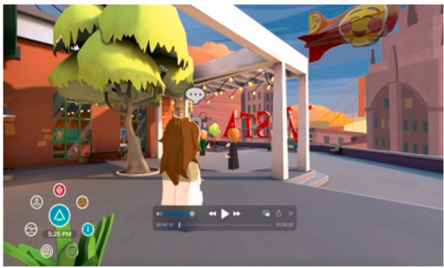
00:24	CF1	CF1 follows the girls.	trails behind (body movement) 	[okay]
-------	-----	------------------------	--	--------

Fig. 2. Multimodal transcript of group 2 (Excerpt 3).

00:32	DF1	DF1 asks a question but does not specify the person the question is addressed. However, we can assume it is CF1 because she is looking at her. She steps a bit back to give CF1 space to go inside, but the moment CF1 moves, DF1 immediately rushes to enter the room first.	uses hand gestures to show the room and rushes first to enter it (body movement) 	<i>"Did you even see this place yet?"</i>
-------	-----	---	--	---

Fig. 3. Multimodal transcript of group 2 (Excerpt 4).

00:37	DF1	DF1 goes in first after asking if CF1 had seen the room before. DF2 follows DF1.	dominates spatially by being the first in the room	
-------	-----	--	--	--

Fig. 4. Multimodal transcript of group 2 (Excerpt 5).

1:06	DF1	DF1 decides on the place where the conversation will take place and initiates the conversation. She also wants to make sure they had read the task and asks the question directly.	positions herself in a different area from her interlocutors (use of space)		<i>"Did you, guys, read the task? I did (.) I don't know where to begin."</i>
------	-----	--	---	--	---

Fig. 5. Multimodal transcript of group 2 (Excerpt 6).

Appendix B

Fig. 6 shows DF1 repeating the same pattern (showing proxemic behaviour) and deciding on the location again (manipulating the VR environment) without waiting for CF1's input. She engages with a variety of modes such as verbal ("*Can we go outside? I like it better there*"), gestural (showing where *there* is) and *spatial* (she is already at the door when she says this), which could potentially create conflict with her Cypriot partner as the latter does not participate again in any decision-making concerning the setting of the task.

Time	Participant(s)	Context	Mode	Language
00:04	DF1	It is the beginning of task 3. DF2 is late so it is only DF1 and CF1 in the living room. Just like in the beginning of task 2, DF1 decides on the place where the conversation will take place and does not give her partner the chance to agree or disagree. She exits the apartment while her partner is still inside.	exits the room (body movement) looks away (face movement) 	<i>"Can we go outside? I like it better there."</i>

Fig. 6. Multimodal transcript of group 2 (Excerpt 7).

00:12	CF1	The two partners are sharing information about snacks they eat at parties and family gatherings. CF1 shares <i>loukoumades</i> , which are traditional honey puffs that people fry. CF1 uses her hand gestures to show the process of frying.	uses hand gestures to illustrate the frying process 	"We have something called <i>loukoumades</i> , which is fluffy small balls, like bread, but they are fried, and they are very sweet."
-------	-----	---	---	---

Fig. 7. Multimodal transcript of group 2 (Excerpt 8).

1:24	CF1	CF1 uploads a picture of <i>loukoumades</i> so that her partner has a visual representation of her traditional snack.	uploads an image of <i>loukoumades</i> to reinforce her point (visual) 	"I can send you a picture of it."
------	-----	---	--	-----------------------------------

Fig. 8. Multimodal transcript of the group 2 (Excerpt 9).

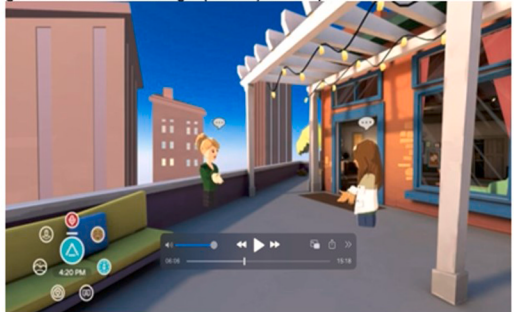
06:06	CF1	Even though DF1 seems disappointed that Mary is not familiar with <i>bitterballen</i> (a Dutch, meat-based snack), she does not make an effort to describe how they look like or how they are made. She immediately asks if CF1 knows of any other typical Dutch snacks. CF1's answer is negative, so DF1 says, "I don't know snacks from your country either". This causes CF1 to laugh awkwardly.	gives an awkward laugh (sound) to her partner's remark 	DF1: "Do you know some other typical snacks from Holland or (.)?" CF1: "I haven't heard of any actually." DF1: "I don't know any snacks from your country either." CF1: [laughs awkwardly]
-------	-----	---	--	---

Fig. 9. Multimodal transcript of the group 2 (Excerpt 10).

Time	Participant	Context	Mode	Language
33:04	DM1	After the completion of the last task, the Dutch student wants to know whether the Cypriot student is interested in visiting the Netherlands.	uses hands (gestures) to emphasise the word <i>three</i> (suggesting they had time to cover a lot of topics) 	"When you think about all the stuff we talked about in these three meetings, would you like to visit the Netherlands?"

Fig. 10. Multimodal transcript of the male dyad (Excerpt 11).

Appendix C

In Excerpt 12 (Fig. 11), the Dutch student immediately recalls that his partner had previously visited the Netherlands. To avoid causing confusion with his repeated question, he explains (“*you already said you did*”) and rephrases it for clarification (“... *but would you like to visit it even more?*”), using strategies that help maintain connection. His explanation is accompanied by laughter and emphatic hand gestures, reinforcing his verbal message. These multimodal cues - higher modal intensity in speech, laughter, and gesture - function as *lower-level actions* that support mutual understanding.

33:12	DM1	After the completion of the last task, the Dutch student wants to know whether the Cypriot student is interested in visiting the Netherlands again .	smiles (facial expression) uses hands (gestures) to emphasize <i>even more</i>		"You already said you did but would you like to visit it even more?"
-------	-----	---	---	--	--

Fig. 11. Multimodal transcript of the male dyad (Excerpt 12).

34:48	DM1	The Dutch student talks about his favourite Dutch food <i>kapsalon</i> .	uses hands (gestures) to show the process of how Dutch people wrap <i>kapsalon</i> with aluminium foil		"You get a (.) you get a kind of (.) how do you call it in English? Like a plate but also in an (.) aluminium]"
-------	-----	--	--	---	---

Fig. 12. Multimodal transcript of the male dyad (Excerpt 13).

34:56	DM1	The Dutch student describes the process of making the Dutch food <i>kapsalon</i> .	uses hands (gestures) to explain the several layers of food		"You get the meat, and French fries, salad and cheese on top of it but it's melted, all together, and then yo, add garlic sauce."
-------	-----	--	---	--	---

Fig. 13. Multimodal transcript of the male dyad (Excerpt 14).

Appendix D. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cexr.2026.100158>.

References

Amadasi, S., & Holliday, A. (2017). Block and thread intercultural narratives and positioning: Conversations with newly arrived postgraduate students. *Language and Intercultural Communication*, 17(3), 254–259.

Arasaratnam-Smith, L. A. (2017). Intercultural competence: An overview. In D. K. Deardorff, & L. A. Arasaratnam-Smith (Eds.), *Intercultural competence in higher education: International approaches, assessment and application* (pp. 7–18). Routledge.

Avgousti, M. I. (2018). Intercultural communicative competence and online exchanges: A systematic review. *Computer Assisted Language Learning*, 31(8), 819–853. <https://doi.org/10.1080/09588221.2018.1455713>

Barbosa, M. W., & Ferreira-Lopes, L. (2023). Emerging trends in telecollaboration and virtual exchange: A bibliometric study. *Educational Review*, 75(3), 558–586. <https://doi-org.proxy.library.uu.nl/10.1080/00131911.2021.1907314>.

Barreda-Ángeles, M., & Hartmann, T. (2022). Psychological benefits of using social virtual reality platforms during the covid-19 pandemic: The role of social and spatial presence. *Computers in Human Behavior*, 127, 1–13. <https://doi.org/10.1016/j.chb.2021.107047>

- Bezemer, J. (2014). 14. Multimodal transcription: A case study. In S. Norris, & C. Maier (Eds.), *Interactions, images and texts: A reader in multimodality* (pp. 155–170). De Gruyter Mouton. <https://doi.org/10.1515/9781614511175.155>.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Byram, M. (2021). *Teaching and assessing intercultural communicative competence: Revisited*. Bristol, Blue Ridge Summit: Multilingual Matters. <https://doi.org/10.21832/9781800410251>
- Calvo-Ferrer, J. R., Melchor-Couto, S., & Jauregi, K. (2016). ReCALL special issue: Multimodal environments in CALL. *ReCALL*, 28(3), 247–252.
- Canto, S., & Jauregi-Ondarra, K. (2017). Language learning effects through the integration of synchronous socializing network opportunities in language curricula: The case of video communication and Second Life. *Language Learning in Higher Education Journal, Special Issue: Synchronous Technologies in Language and Intercultural Learning and Teaching in Higher Education*, 7(1), 21–53. <https://doi.org/10.1515/cercles-2017-0004>.
- Canto, S., Jauregi-Ondarra, K., & Gruber, A. (2023). High-Immersion Virtual Reality-Mediated Intercultural Virtual Exchange: A Case Study. In S. Di Sarno-García, S. Montaner-Villalba, & A. Gimeno-Sanz (Eds.), *Telecollaboration Applications in Foreign Language Classrooms* (pp. 136–159). IGI Global. <https://doi.org/10.4018/978-1-6684-7080-0.ch007>.
- Chen, H. L., & Sevilla-Pavón, A. (2023). Negotiation of meaning via virtual exchange in immersive virtual reality environments. *Language, Learning and Technology*, 27(2), 118–154. <https://hdl.handle.net/10125/73506>.
- Chimeva, Y., & Trenchs-Parera, M. (2024). Different interlocutors, different EFL interactional strategies: A case study of intercultural telecollaborative projects in secondary classrooms. *ReCALL*, 36(1), 104–118. <https://doi.org/10.1017/S0958344023000228>
- Çiftçi, E. Y., & Savaş, P. (2018). The role of telecollaboration in language and intercultural learning: A synthesis of studies published between 2010 and 2015. *ReCALL*, 30(3), 278–298. <https://doi.org/10.1017/S0958344017000313>
- Cunningham, D. J., & Vyatkina, N. (2012). Telecollaboration for professional purposes: Towards developing a formal register in the foreign language classroom. *Canadian Modern Language Review*, 68(4), 422–450.
- Deardorff, D. K. (2006). The identification and assessment of intercultural competence as a student outcome of internationalization at institutions of higher education in the United States. *Journal of Studies in International Education*, 10(3), 241–266.
- Deardorff, D. K. (2009). Synthesizing conceptualisations of intercultural competence: A summary of emerging themes. In D. K. Deardorff (Ed.), *The SAGE handbook of intercultural competence* (pp. 264–269). Thousand Oaks, CA: Sage.
- Dervin, F. (2016). Tools for change—dynamic and realistic intercultural competences. In *Interculturality in education*. London: Palgrave Pivot. https://doi.org/10.1057/978-1-137-54544-2_6.
- Dhimolea, T. K., Kaplan-Rakowski, R., & Lin, L. (2022). A systematic review of research on high-immersion virtual reality for language learning. *TechTrends*, 66, 810–824. <https://doi.org/10.1007/s11528-022-00717-w>
- Duff, P. A. (2020). Case study research: Making language learning complexities visible. In T. McNamara, & L. Plonsky (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 144–153). Routledge, 2020.
- Ellis, R. (2003). *Task-Based Language learning and teaching*. Oxford University Press.
- Fantini, A. E. (2020). Reconceptualizing intercultural communicative competence: A multinational perspective. *Research in Comparative and International Education*, 15(1), 52–61. <https://doi.org/10.1177/1745499920901948>
- Freina, L., & Ott, M. (2015). A literature review on immersive virtual reality in education: State of the art and perspectives. *The international scientific conference eLearning and software for education*, 1(133), 10–1007.
- Gruber, A., Canto, S., & Jauregi-Ondarra, K. (2023). Exploring the use of social virtual reality for virtual exchange. *ReCALL*, 35(3), 258–273. <https://doi.org/10.1017/S0958344023000125>
- Gruber, A., & Kaplan-Rakowski, R. (2020). User experience of public speaking practice in virtual reality. In R. Zheng (Ed.), *Cognitive and affective perspectives on immersive technology* (pp. 235–249). IGI Global. <https://doi.org/10.4018/978-1-7998-3250-8.ch012>.
- Guichon, N., & Cohen, C. (2016). Multimodality and CALL. In F. Farr, & L. Murray (Eds.), *The Routledge handbook of language learning and technology*. Routledge.
- Helm, F., & Dooley, M. (2017). Challenges in transcribing multimodal data: A case study. *Language, Learning and Technology*, 21(1), 166–185, 10125/44600.
- Holliday, A. (2016). Difference and awareness in cultural travel: Negotiating blocks and threads. *Language and Intercultural Communication*, 16(3), 318–331.
- Hwang, Y., & Lee, S.-M. (2024). ‘Can we go to Tower Bridge?’: Teaching British culture from text to immersive realities. *Innovation in Language Learning and Teaching*, 1–12. <https://doi.org/10.1080/17501229.2024.2439406>
- Jauregi-Ondarra, K., Meijerink, J., & Christoforou, M. (2024). Using high-immersion social virtual reality environments for researching interculturality. In K. Sadeghi (Ed.), *Routledge Handbook of Technological Advances in Researching Language Learning* (pp. 403–419). Routledge. <https://doi.org/10.4324/9781003459088-36>.
- Jewitt, C. (Ed.). (2014). *The routledge handbook of multimodal analysis* (2nd ed.). Routledge.
- Karimi, H., Sañosa, D. J., Rios, K. H., Tran, P., Chun, D. M., Wang, R., & Arya, D. J. (2023). Building a city in the sky: Multiliteracies in immersive virtual reality. *CALICO Journal*, 40(1), 24–44.
- Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT press.
- Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse – The modes and media of contemporary communication*. London: Arnold.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press.
- Lee, S. M., Wang, X., Park, I., & Lestiono, R. (2024). It feels so real! situated authentic language learning in immersive virtual reality. *Education and Information Technologies*, 1–23.
- Lewis, T., & O’Dowd, R. (2016). *Online intercultural exchange and foreign language learning: A systematic review* (pp. 21–66). Online intercultural exchange.
- Liaw, M. L., Priego, S., Dieng, D., & Doyon-Lacasse, K. (2024). The agency development of teachers of an additional language via immersive VR telecollaboration. *Computer Assisted Language Learning*, 1–32. <https://doi.org/10.1080/09588221.2024.2350415>
- Lin, V., Barrett, N. E., Liu, G. Z., Chen, N. S., & Jong, M. S. Y. (2023). Supporting dyadic learning of English for tourism purposes with scenery-based virtual reality. *Computer Assisted Language Learning*, 36(5–6), 906–942.
- Meijerink, J. (2023). *Negotiating interculturality in high-immersive VR: A case study analysis on intercultural communication during a virtual exchange*. Utrecht University, Master thesis.
- Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, 77(12), 1321–1329.
- Mills, K. A., Scholes, L., & Brown, A. (2022). Virtual reality and embodiment in multimodal meaning making. *Written Communication*, 39(3), 335–369. <https://doi.org/10.1177/07410883221083517>
- Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. Routledge.
- Norris, S. (2016). Concepts in multimodal discourse analysis with examples from video conferencing. *Yearbook of the Poznan Linguistic Meeting*, 2(1), 141–165.
- Norris, S. (2019). *Systematically working with multimodal data: Research methods in multimodal discourse analysis*. Wiley.
- O’Dowd, R., & Dooley, M. (2020). Intercultural communicative competence development through telecollaboration and virtual exchange. In *The Routledge handbook of language and intercultural communication* (pp. 361–375). Routledge.
- Palomeque, C., & Pujolà, J.-T. (2018). Managing multimodal data in virtual world research for language learning. *ReCALL*, 30(2), 177–195. <https://doi.org/10.1017/S0958344017000374>
- Pirini, J., Matelau-Doherty, T., & Norris, S. (2018). Multimodal analysis. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The palgrave handbook of applied linguistics research methodology*. London: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-59900-1_28.
- Satar, M. (2020). L1 for social presence in videoconferencing: A social semiotic account. *Language, Learning and Technology*, 24(1), 129–153, 10125/44713.
- Schenker, T. (2012). Intercultural competence and cultural learning through telecollaboration. *CALICO Journal*, 29(3), 449–470.
- Tan, S., O’Halloran, K. L., & Wignell, P. (2016). Multimodal research: Addressing the complexity of multimodal environments and the challenges for CALL. *ReCALL*, 28(3), 253–273. <https://doi.org/10.1017/S0958344016000124>
- Wigham, C. R., & Satar, M. (2021). Multimodal (inter)action analysis of task instructions in language teaching via videoconferencing: A case study. *ReCALL FirstView*, 1–19. <https://doi.org/10.1017/S0958344021000070>
- Zhang, X., & Zhou, M. (2023). Information and digital technology-assisted interventions to improve intercultural competence: A meta-analytical review. *Computers & Education*, 194. <https://doi.org/10.1016/j.compedu.2022.104697>