

Research Article

Timo Ahlers*, Cassandra Bumann, Ralph Kölle, and Milica Lazović

Foreign Language Tandem Learning in Social VR

Conception, Implementation and Evaluation of the Game-Based Application *Hololingo!*

<https://doi.org/10.1515/icom-2021-0039>

Abstract: *Hololingo!* is a social virtual reality app for real-time immersive distance learning of German as a Foreign Language (GFL). The acquisition of discursive oral language skills for applied authentic contexts is challenging for group-based classroom settings and is often outsourced to autonomous analogue tandem learning. We operationalise a Digital Game-Based Language Learning (DGBLL) approach for distance learning, which relieves tandems from overstraining autonomy and self-guidance. *Hololingo!* supports tandems in their learning activities by providing entertaining communicative, collaborative, and didactically designed team tasks. These are embedded in a narrative adventure, target linguistic phenomena, and support the holistic acquisition of oral discourse competencies and fluency. The combination of the tandem principle, immersive Social VR gamification, barrier-free access to a global pool of expert/native speakers, and a curricular connection through task selection facilitates language learning and provides transcultural contacts. App development follows a Design-Based-Research cycle of conception, implementation, and evaluation. The results of a qualitative discourse analysis of the learning interaction are discussed with regard to tandem role behaviour, the relation between task design and communication as well as affective and creative behaviour.

Keywords: Social Virtual Reality, Language Tandem, Digital Game-Based Language Learning

1 Introduction: The *Hololingo!*¹ Project

In a mobile, global world, effective foreign and second language learning plays a key role in migration societies to foster participation and integration. Arrivers need language skills for access, participation, and integration regarding communities, companies, and societies, which need to access newcomers' skills and knowledge. However, learning a new language is not trivial and often takes years to master. Therefore, easy access to learning material and swift integration of learning opportunities into everyday activities is crucial for rapid success. Especially opportunities of learning the language in authentic contexts with expert/native speakers of the target language are rare. This is especially the case when the target language is learnt in a foreign country, where it is not established. An international survey among learners of German as Foreign Language (GFL) showed that their biggest goal, speaking the target language fluently, was considered at the same time by far their weakest competence [2]: On average, GFL-learners rated their productive oral-discursive language skills one competence level² lower than their receptive skills, and their speaking skills a half competence level lower than their writing skills [ibid.]. Of course, the spontaneous speaking of a language in authentic applied contexts with expert/native speakers is a challenging task environment that requires fast lexical access and broad lexical knowledge, confident command of pronunciation, grammar, interactional and discursive-pragmatic competencies (making propositions, arguing, turn taking), planning of utterances while listening, and interpretation of everyday language. However, the weakness of spoken language skills also results from too little practice. Oral-discursive competencies are hard to target by i) group-based teaching settings, where learners usually communicate among themselves, and by ii) behaviorist text in-

*Corresponding author: Timo Ahlers, Universität Potsdam, Potsdam, Germany, e-mail: timo.ahlers@uni-potsdam.de, ORCID: <https://orcid.org/0000-0001-7565-4306>

Cassandra Bumann, Ralph Kölle, Universität Hildesheim, Hildesheim, Germany, e-mails: bumann@uni-hildesheim.de, ralph.koelle@uni-hildesheim.de

Milica Lazović, Philipps-Universität Marburg, Marburg, Germany, e-mail: milica.lazovic@uni-marburg.de

¹ The article is an extended version of [1].

² A1, A2, B1, B2, C1, C2 according to the Common European Framework of Reference for Languages [11].

tense language learning apps (*Babbel*, *Duolingo*, *MondlyVR*). Both language learning methods also occur in non-applied contexts in imagined “as if” scenarios (e. g. learners pretend to order a coffee or make small talk). Here, the learners are always in training mode with very little authentic feedback on the effectiveness and appropriateness of the practised verbal actions. The same problem applies to new social VR-based scenarios that transfer language teaching from an analogue to a virtual classroom³ or conduct group-based field trips, e. g. virtual library tours:⁴ For organisational reasons, participants get relatively little individual speaking practice and do not learn as much as they could by talking to peer learners of the same language level. Accordingly, learners frequently cite difficulties in oral communication in the target language: lack of fluency, problems in understanding varieties/everyday language and suffer from speech anxiety [2]. Immersive, situative learning by language travel or exchange programs is considered highly effective, but can be expensive, time-consuming, and hard to integrate into everyday life. Therefore, language practice is often outsourced to private, autonomous tandem learning [6], where students exchange their expert/native languages in self-directed analogous settings, e. g. helping each other with homework or simply hanging out together. In a previous study, German learners indicated tandems as the most popular method [2].

Nevertheless, the analogue tandem method has clear limits: a) a restricted number of suitable learning partners on-site (incongruous availability, language match, sympathy), b) an effort of coordinating and going to meetings), c) overtraining self-directed learning (choice of assignments, topics, methods) and d) a lack of tracking the learning progress [9, 2]. As a result, analogue tandems are compared to courses and apps only little spread and often quickly disused. Currently, digital video tandem apps (e. g. *Hellotalk*, *Tandem*) are popular, as indicated by increasing download numbers. They provide easy access to a global pool of tandem partners through mostly already owned mobile devices like smartphones. However, digital tandems [14] neither substantially support self-directed learning nor help learners to keep track of their learning progress. Also, video tandems quickly generate the typical

and boring ‘talking, to practice talking’ due to a lack of entertaining joint activities. Video tandem partners permanently need to come up with topics, questions, and assignments themselves. They must fully construct their learning process, which can be overtraining [2, 19].

Our Social Virtual Reality (SVR) approach masters these problems: Through the virtual tandem method, the application *Hololingo!* not only potentially enables location-independent, time-flexible language learning and low-threshold access to a global user pool. It also offers activating collaborative, didactically designed tandem tasks for Digital game-based language learning (DGBLL, [18]) connected to a narrative, immersive SVR adventure. The task environment provides enabling opportunities for communicative, empractic, collaborative practice and spoken language learning in an immersive 3D environment. DGBLL-tasks can target i) specific language phenomena – and thus deliver the possibility of curricular connection and integration – and ii) holistic oral-discourse practice in applied contexts with native/expert speakers in which tandem partners are aware of their respective roles. Learners shall lead conversations and make meaningful, comprehensive contributions while expert/native speakers are aware of the learners’ needs and shape a supportive setting for them. Expert speakers are prepared to act as linguistic role models and to adjust their articulation and utterances to the learners’ needs through awareness, patience, interest, and communicative grounding. They shape a supportive and scaffolded setting, in which they give learners space, time, help (e. g. vocabulary offers), feedback and encouragement to put their thoughts into complex language [19], but also acknowledge them as equal game partners (facilitated by collaborative tasks), despite the hierarchy in language competence. The immersive SVR-DGBLL task environment transfers imagined learning scenarios of classroom settings into experienced motor-stimulating empractic language scenarios, making applied language skills easier to learn and to transfer to analogue contexts [3]. Yet, the success of VR apps is tied to the increasing use of mass-market hardware. The app-project is carried out by the interdisciplinary workgroup *Foreign Language Learning in VR* by the Universities of Potsdam, Hildesheim, and Marburg⁵ in a Design-Based-Research manner [20] with iterative development cycles of (re-)conception, implementation, and evaluation.

³ Offer on vr language courses in a classroom setting in Rec Room: Video “LANGUAGE EXCHANGE IN VR – You Might Even Find Your Perfect Mate! Learn Languages During Lockdown!”. 2020, by MRTV-MIXED REALITY TV, https://youtu.be/tx8rKy1_OAc (26.01.2022).

⁴ Video on a language learning field trip to a virtual library in AltspaceVR: “Learn English in VR in a Library in Altspace”. 2020, by VR English Lessons, <https://youtu.be/Y4sKxRlFloo> (26.01.2022).

⁵ Website: <https://hololingo.blog.uni-hildesheim.de/> (26.01.2022).

2 Related Work and Concepts

Digital Game-Based Language Learning (DGBLL): “[A] learning game is defined as a playful activity that is structured by rules for the pursuit of quantifiable outcomes (e.g. win states and points), and incorporates educational objectives (e.g. knowledge acquisition) as its own end” [18]. DGBLL comprises educational language learning games for first and second/foreign language acquisition. It is considered a highly beneficial method, because of “immersive exposure to the language learning environment, lowered anxiety and other affective barriers to language learning, and increased use of the target language for interaction in gaming” [18]. The field of DGBLL has a particular research focus on language acquisition in digital multiplayer role-playing games like *Second Life* [7] or *World of Warcraft* “where language learners interact and communicate for authentic purposes in 3D virtual worlds” [18]. A comprehensive review [18] reports that DGBLL delivers better results than traditional language learning in many aspects: higher learning duration, motivation, experienced self-efficacy, the will to communicate in the target language, engagement in written communication outside the game (forums, private chats), pragmatic acquisition of appropriate language use (politeness, humour) and cultural learning.

Social Virtual Reality Tandem Learning: Like successful applications for vocational training [26], *Hololingo!* uses Virtual Reality, which increases physical 3D immersion by providing i) an ego perspective, ii) sensorimotor coupling to an avatar, and iii) interactive worlds through stereoscopic head-mounted displays, motion, and hand controller tracking and 3D sound. While single-player VR language learning applications are limited to semi-authentic, scripted communication scenarios with chatbots (e.g. *Mondly VR*) or to vocabulary learning (*Word Saber*, [16]). Multiplayer apps (e.g. *AltSpaceVR*, *VRChat*, *Rec Room*) provide public and private chat rooms with entertaining hangout activities (bowling, dancing, snowball fight). In these environments, some lecturers offer group-based language courses in classroom or field trip settings [19]. But Social VR apps can also be used for autonomous tandem learning, where two native speakers of different languages support each other in learning the other’s language. Based on mutuality and autonomy, tandem learning benefits from direct contact with expert/native speakers to correct and consolidate foreign language skills in practice [6]. An analysis of 1:1 tandem communication in *AltSpaceVR* [2] found complex multimodal interactions (e.g. learning the integrated use of

deictics and manual pointers), indicating transferability to analogue contexts. However, virtual and analogue tandems exhibit potentially overburdening learning autonomy, which manifests in a lack of control over the learning process and is expressed in the desire for more feedback [2]. Feedback and reflection are also often missing in conventional single-user smartphone apps [17]. For our goal to relieve Hangout tandems of overwhelming self-direction and still evoke essentially free, entertaining, and learning-focused tandem discourse, *Hololingo!* combines the expert-novice tandem setting with cooperative, didactically designed DGBLL-tasks and a captivating storyline.

Design-based research (DBR): For app development, we use a DBR approach [20]. After conceptualising and implementing the first demonstrator, we initiated an iterative development process of analysis/exploration, design/construction, and evaluation/reflection. First, we carried out formative qualitative tests with German speakers to optimise usability. Then, we performed a first user test with a tandem. We recorded and linguistically analysed audio and video data (VR ego perspective, analogue scene) and examined the results for correspondences between task design and elicited communication to improve the theoretical understanding and optimise the demonstrator. We also collected both tandem partners’ oral and written user feedback regarding the tandem experience and app usability, which we will use to develop and design a prototype.

3 Didactic Conception

Hololingo! is an app project in progress. Our overall goal is to develop a DGBLL-SVR app for tandem learning that users can access from all over the world. Learners will create profiles with languages, competence levels, and interests to match with other users. Based on curricular proximity to their competence level according to the Common European Framework of Reference for Languages (CEFR) [11], respective learning scenarios will be offered for i) the development of holistic oral-discursive language competencies [8] and ii) the practice of specific linguistic (lexical, grammatical, and pragmatic) phenomena. In future upgrades, learners shall keep track of their progression by tests, coupled to a badge system. For now, the demonstrator comprises a first holistic learning scenario, *Myth of the Huckup*. A statue of Huckup, a mythical troll, is located in Hildesheim’s city centre (Figure 1). As a metaphor for guilt, it jumps down from a tree in the necks of apple thieves



Figure 1: Statue of *Huckup*⁶ in the city centre of Hildesheim, inscription⁷, and the digital model of *Huckup* in *Hololingo*!

for punishment. We created a thematic adventure where tandems learn about the myth and relive it through various DGBLL exit-game stations (Figure 2).

First, the tandem partners get to know each other, watch an introductory video on tandem learning⁸ and negotiate their roles for the upcoming tandem work (station 1). They then learn about the myth by mutually reading an introduction from stone steles (st. 2). They further engage with the saga by translating the inscription of the Huckup statue from a Low German text into standard German – a task that can only be solved by collaborative linguistic reflection (st. 3). Afterwards, the exit-room game starts. Users must verbally coordinate two distant switches' flipping to open the door to the next room. Here, they further connect to the myth via a reflection task on a spooky short

film.⁹ In a subsequent collaborative word puzzle, they relive the saga by stealing apples: First, they must place the keyword with apples on correct letter-tiles (st. 4). Then they must find their path through the maze and steal more apples on their way, which they need to put on a table in the break room to unlock the next door (st. 5). In the break room, tandem partners can relax – even take a VR break and put down their HMDs for a moment – or have some hangout and small talk time together (st. 6). In the second part of the maze, they must escape from the troll (if else, they are relocated to the beginning). They have to watch out, warn each other and navigate collaboratively through the maze (st. 7). After the escape, there is another opportunity for a short break before engaging in a reflection task: securing understanding of the myth, feedback on the language learning process (st. 8). In Huckup's cabin, it follows a spontaneous narration and exchange about a comparable myth from the learner's cultural sphere and a reflection on the joint tandem practice (st. 9). Finally, tandem partners may decide to stay in contact and say goodbye to each other (st. 10).

Our main goals are i) to increase users' awareness of their respective tandem role as learner or expert, and

⁶ Hildesheim-Hoher.Weg.Huckup.01., by Longbow4u, 8 July 2005, photograph, 2.136 × 2.848, Wikimedia Commons, public domain.

⁷ Detail of: Huckup, by Ramessos, 11 January 2008, photograph, 948 × 1.230, Wikimedia Commons, public domain. Eastphalian dialect inscription: Junge lat dei Appels stahn / Süs packet deck dei Huckup an. / Dei Huckup is en starken Wicht. / Hölt mit dei Stehldeifs bös Gericht. [„Boy, let go of the appels / or the ‚Jump-on‘ will catch you / the ‚Jump-on‘ is a strong troll / who punishes thieves“].

⁸ Video on tandem learning: Sprachen lernen – im Tandem an der Leuphana. 2014. Leuphana Universität Lüneburg. <https://youtu.be/XcSiS4DWmpY> (26.01.2022).

⁹ Short film: Huckup. Die Hildesheimer Sage vom Huckup, filmisch umgesetzt [The Hildesheim legend of Huckup]. 2018. Flimmerfabrik, <https://youtu.be/VV5agimByD0> (26.01.2022).

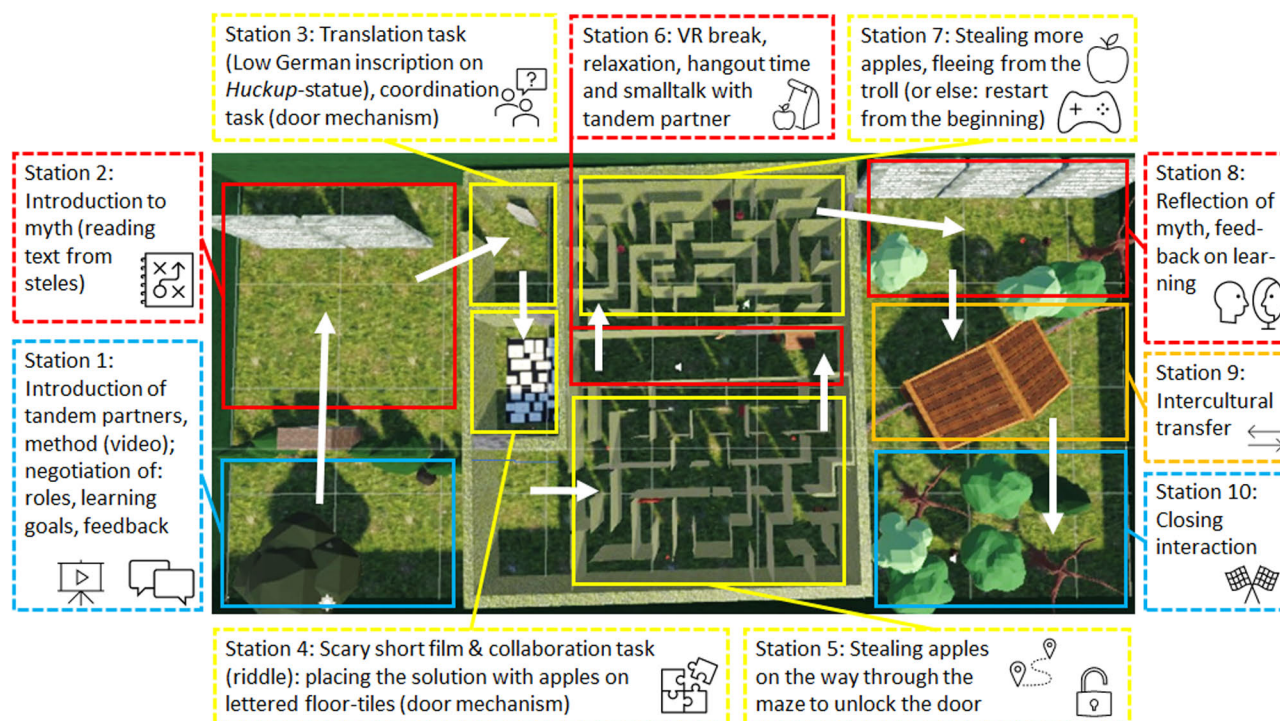


Figure 2: Plan of the *Unity*-based *Hololingo!* adventure from an aerial perspective.

ii) to evoke as much learning-enhancing, communicative language practice as possible by the game-based, collaborative-communicative task design. In the implementation of these goals, different learning competence levels can be targeted, as our scenario shows concerning the learning goal taxonomy of Anderson & Krathwohl [5]: The Hildesheim legend of Huckup is introduced through a receptive text and video task (stations 2 and 4), which correspond to the competence levels of i) remembering and ii) understanding of e.g. vocabulary and content. Other tasks trigger problem-solving competences of iii) application, iv) analysis, v) reflection: The word puzzles (stations 4 and 7) and the cultural transfer task (station 9) can be seen as applying vocabulary and knowledge to master game relevant goals. The translation task (station 3) requires a comparative analysis. The free and the guided reflection phase (stations 6 and 8) facilitate evaluation competence. The iv) creation competence is addressed by collaborative-coordinative exit room tasks, communicative planning, and joint-actional execution of solutions for the door mechanisms. Thus, VR tandem scenarios can be used both a) as a free learning opportunity for tandems and b) to assess previously taught competencies by teachers.

4 Implementation

We used *HTC VIVE* and *HTC VIVE Pro* head-mounted displays (HMDs) and recently also the *Oculus (Meta) Quest 2*, a stand-alone HMD that does not require external tracking stations, like other mass-market and console-based solutions. VR devices allow users to immerse themselves in virtual environments through stereoscopy, motion, and hand tracking [26]. We did not yet use further early-stage hardware additions, like body, finger, eye, or face tracking, which exist for the *VIVE* platform but plan to integrate them eventually, as they might increase the immersive user experience and open up new possibilities for analysing behaviour. Social VR apps gather spatially separated people in a joint virtual environment. Users are represented by avatars and experience the virtual surroundings from the ego perspective. Although there are already some social VR platforms (e.g. *AltspaceVR*, *VRChat*, *Rec Room*), VR learning spaces that specifically evoke linguistic phenomena and boost holistic language learning for freely interacting tandem teams do not yet exist. With *Hololingo!*, we created a first didactically designed DGBLL environment for SVR. Based on a user-centred design approach, first iterations of the *Hololingo!* world have been developed using *Unity*, the *VRChat* SDK (Software Development Kit).

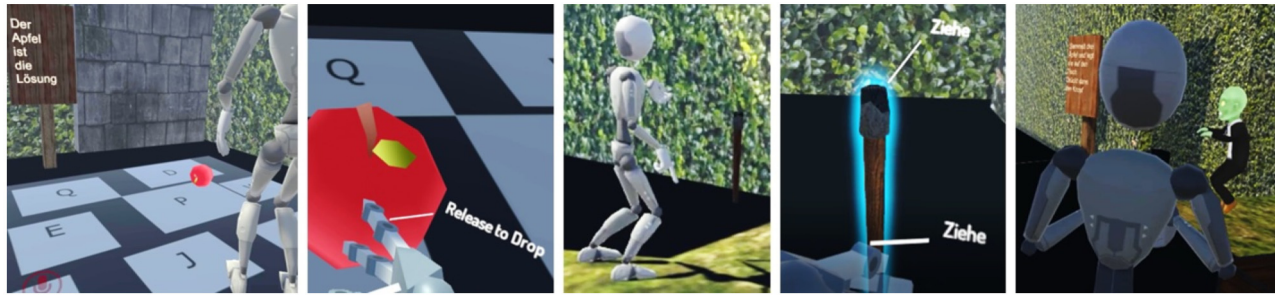


Figure 3: Word puzzle with apples (Figure 2, st. 4), door mechanism: simultaneous lever pull (st. 3), escape task with troll (st. 7).

opment Kit) and the *HTC VIVE (Pro)* HMDs. *Unity* is a cross-platform manufacturer-independent game engine, while the *VRChat* SDK provides basic methods for building social rooms with body and movement synchronisation as well as voice chat, which is crucial to support language learning. In our experience, developing a world with moderate interaction capabilities was straightforward, and low coding experience is required, but developers should be comfortable using the *Unity* engine. The *VRChat* SDK provides a range of triggers and actions to implement interaction. We added functions like pressing buttons and interacting with objects by grabbing or pulling on them to perform, for example, an exit-room coordination task in which users had to synchronise verbally on the simultaneous operation of two distant levers to unlock a door (Figure 3, middle). We integrated videos to introduce the tandem method and as conversation starters. For complex interactions, like spelling a word by dropping apples on lettered floor tiles (Figure 3, left), we used animations and triggers from *Unity* and actions provided by *VRChat*. We also developed gaming elements like escaping from Huckup in a maze – a *Pacman* scenario where users must verbally coordinate for navigation and warnings (Figure 3, right).

As mentioned earlier, the educational design process is related to the Design-Based-Research method [20] by iteratively analysing the problem domain, designing and evaluating a solution, formatively and summatively. The demonstrator was developed between May and July of 2019 and is a private world on *VRChat*. However, there are further options to develop virtual environments based on Software Development Kits (SDKs) from HMD manufacturers. Recently, more cross-manufacturer platforms came up (*OpenVR*, *Unity XR Interaction Toolkit*). *VRChat* is a SVR platform that allows users to interact in virtual worlds and develop their own. In the beginning, we chose the *VRChat* SDK because of a low-threshold entry into programming as well as a rich set of objects and features, visually appealing environments, and refined avatars.

In contrast, the SDK does not support the implementation of complex interaction capabilities well. This led to various issues, like animations getting stuck at different conditions and not executing correctly. This impacted the experience of test users negatively for some tasks. Other downsides of using the SDK were an elaborate installation and the dependence on the *VRChat* platform itself. Even though using the SDK resulted in an appealing learning world, it was not (yet) possible to publish the demonstrator as a public world on *VRChat* or in an open-access format to make it available to other researchers for subsequent use or development.

Due to these limitations, the demonstrator [10] has been migrated to a self-developed, cost-effective solution using *Unity* again, replacing the *VRChat* SDK with *Photon Unity Networking (PUN)*¹⁰ and *Photon Unity Voice*. *PUN* is a *Unity* asset to realise multiplayer games by providing authentication options, matchmaking and in-game communication through the *Photon* server backend. The *PUN* multiplayer features are based around the lobby and social VR room creation. Software is hosted in a globally distributed *Photon* cloud environment to guarantee low latencies for players worldwide. *PUN* exports to almost all platforms supported by *Unity*, which is important for the further *Hololingo!* development. There are two different packages to choose from: *PUN Free*, which is free up to 20 concurrent users (CCU) and *PUN Plus* for more than 20 CCU. We integrated *Unity XR Interaction Toolkit* as a second new component, a high-level, component-based interaction system. It provides a framework that makes 3D and UI interactions available from *Unity* input events. *Unity XR Interaction Toolkit* contains a set of components that support interaction tasks like cross-platform XR controller input, basic object hover, select and grab, haptic feedback through XR controllers, visual feedback (tint/line rendering) to indicate possible and active interactions and a VR

¹⁰ Website: <https://www.photonengine.com/pun> (26.01.2022).

camera rig for handling stationary and room-scale VR experiences. The *XR Interaction Toolkit* is in a very dynamic state of development,¹¹ which makes it, on the one hand, difficult to keep projects up-to-date; on the other hand, it abstracts interactions in VR for all relevant VR platforms. The project currently supports the *HTC Vive (Pro)* as well as the *Oculus (Meta) Quest (2)*. Based on the independence from *VRChat* and regarding the *HTC Vive Pro-Eye* with its integrated eye tracker, we plan to perform eye tracking studies to retrieve sophisticated data and learn more about the users' interactions within the *Hololingo!* world.

5 Evaluation

By the following research questions, we examined core design elements of the current *Hololingo!* demonstrator: Does the design i) enhance L2 learners' discourse participation, ii) contribute to the performance of different tandem roles?,¹² iii) create opportunities for intense conversation, iv) entertain and generate fun? We compared hangout to DGBLL tandems and conducted a qualitative communication analysis.

5.1 Hangout vs. DGBLL Tandem Communication

In a previous study, we evaluated 3 h 55 min. of audio and video data from 13 self-directed hangout tandem conversations in *AltspaceVR*, recorded with exchange students in 2019 at the University of Hildesheim. We also analysed 1 h 27 min. of DGBLL tandem data, reflecting the latest version of the *Hololingo!* app [19].¹³ Each tandem consisted of a L1 expert and a L2 learner of German communicating from different lab rooms via *HTC VIVE Pro* devices. In both settings, participants were informed about the tandem method and got a short VR training before meeting their tandem partners in VR. In the didactic DGBLL setting,

tandem partners also watched a short educational video on tandem learning in VR together and got the task to reflect on it (Figure 2, st. 1). A quantitative analysis identified the word count for L1 experts and L2 learners for both data sets, including discourse particles (e. g. *ähm, mh*). The results showed a similar word/minute ratio for hangout vs. DGBLL tandems but a significant difference for discourse participation [ibid.]. In the self-directed hangout data of *AltspaceVR*, L2 learners accounted for only 32 % of the words, while for 46 % in the *Hololingo!* data. The didactic design seems to enhance learner participation [ibid.]. However, a more controlled follow-up study with more participants than the examined one DGBLL tandem needs to test the results' validity.

5.2 Qualitative Analysis of DGBLL Tandem Discourse

We examined the DGBLL *Hololingo!* data of a Chinese exchange student learning German (L2; level B2) and a German native speaker (L1), who had not met before. We transcribed selected parts of the 87 min recording in the HIAT format and carried out a multimodal functional-pragmatic communication analysis [13]. A previous analysis of the translation and reflection tasks (Figure 2, st. 1 and 6) shows high degrees of joint communicative interaction (e. g. co-constructions to ground mutual understanding, multimodal use of gestures and emojis) [19]. Now, we focus on i) tandem role typical behaviour for expert and learner, ii) elicitation of communication by task design, and iii) the affective learning experience and indications of fun.

5.2.1 Tandem Role: L1 Expert

The tandem setting was only briefly explained to the participants before the recording. After greeting each other in VR, they were informed about the tandem method by a video and a reflection task (Figure 2, st. 1). We observed largely role-specific behaviour for the L1 expert and the L2 learner.

L1 adjusts in speed, volume, and clarity of pronunciation to L2's needs by communicative grounding. L1 takes control of the interaction in critical phases and guides L2 (L1: *Bevor wir das Video abspielen, steht da [...], dass wir uns einander vorstellen müssen* [Before we play the video, it says (...) that we must introduce ourselves to each other]). L1 also uses various strategies to give L2 the turn and opportunities to speak: Pauses show that L1 patiently waits

¹¹ Website: <https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@2.0/manual/index.html> (26.01.2022).

¹² We will implement DGBLL tasks for elicitation of *specific* language phenomena in a later demo, as they had been proven successful in other contexts [25, 4]. Our focus is on *holistic* tasks for the current version.

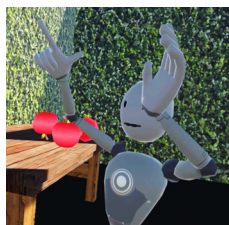
¹³ We thank all participants, all students of the seminars *Interkulturelle Kommunikation/Fremdsprachenlernen in SVR*, and M. Goecht, A. Schwanke, K. Köller, F. Breker, A. Taranto, S. Elbeshausen, K. Schweiger and K. Senkbeil for their contribution and support.

for answers, reactions, or initiatives from L2 after own contributions (*Wir müssen uns erst vorstellen!* [10 sec.] *Hallo?* [We must introduce ourselves first! (10 sec.) Hello?]). L1 lets L2 go first several times (L2: *Also, ich fange an, oder?* L1: *Ja.* [L2: Well, I will start, shall I? L1: Yes]. L1 encourages L2 to make contributions by asking her direct questions (*Warum hattest du gelacht, als, als der eine ...* [Why did you laugh when, when the one...?]). L1 also hands the turn back by follow-up questions to invite L2 for further elaboration (L1: *Er hat gesagt, Chinesen singen gern Karaoke?* L2: *Jaa.* L1: *Okeh...?* [L1: He said Chinese people like to sing karaoke? L2: Yes. L1: Okay...?]). L1 rarely leaves out conversation opportunities, e. g. by rushing to the next task (L2 answers L1's question. L1: *Ok, also. Die dritte Aufgabe – durchs Tor gehen* [L1: Ok, so. The third task: go through the gate]). Instead, L1 sometimes follows up on L2 persistently to make sure that L2 has understood the content: L1 uses *teacher questions*, to which she already knows the answer, to get L2 to verbalise her thoughts (L1: *Was haben wir gelernt?* [What have we learnt?]). When L2 asks about content, L1 answers and states her own understanding. Overall, L1 acts as an attentive, helpful partner who invites L2 to speak a lot and gives helpful feedback (L2: *Wie heißt das? Er stiehlt? Stiehlt?* L1: *Stiehlt!* [L2: What is the form? He steal? Steals? L1: Steals!]). She gives necessary feedback also when it is not requested (L2: *Er stahl den Glocke.* L1: *DIE Glocke* [L2: He stole the bell (wrong gender). L1: THE bell (correct form)]).

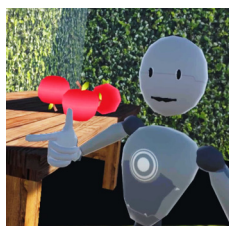
5.2.2 Tandem Role: L2 Learner

L2 also performs her role as a learner well. When she does not know the vocabulary, she gives verbal and gestural explanations until L1 names the word.

L2: *Wie heißt das? Äh, ein RING oder so? [...]* so wie DAS. [What is it called? Uh, a RING or something? (...) so yeah, like THAT.] (outlines a bell shape with her hands)



L2: *und es klingt, klingelt* [and it sounds, rings] (repeatedly strikes the previously drawn bell with her right index finger)



L2 explicitly requests feedback on grammatical forms, when she is not confident (L2: *Wenn es in Vergangenheit, also er STAHL?* L1: *Jaaa* [L2: When in the past tense, so he STOLE? L1: Yes.]). She sometimes offers varying forms until L1 confirms or states the correct form (L2: *Äh, die – DER böse Wicht?* L1: *Mmh!* [L2: Uh, the (incorrect gender), THE (correct form) evil troll? L1: Mmh!]; L2: *... auf dem Rücke?* *Auf dem rück?* L1: *RückEN?* [L2: ... on his back? (2x with wrong inflection) L1: Back? (correct form)]). L2 eventually takes more initiative by reading out assignments and suggesting solutions first: (Both approaching station 3) L2: (reads out loud the hint) *Ihr kommt nicht weiter, wenn ihr nicht zusammen arbeitet.* L1: *Ok.* L2: *Und, ich glaube es hat mit diesem Steintür zu tun, oder?* [L2: You will not get anywhere if you do not work together. L1: Ok. L2: And, I think it has to do with this stone door, hasn't it?]). L2 effectively uses opportunities that are provided to her by L1 for practice and learning. She seems to appreciate L1's support by uttering interest in talking to L1 after the game (L2: *Ich kann dir erzählen, vielleicht später?* [L2: Maybe I can tell you later about it?]).

5.2.3 Elicitation of Communication Through Task Design

For *Hololino!* it is essential to design DGBLL tasks that reliably elicit appropriate communication for language practice and learning. Unlike topic-based conversation and reflection tasks (Figure 2, st. 2, 6, 8, 9), we also aimed for a) collaborative tasks of sharing problem understanding and constructing solutions that address interactional competencies with frequent turn taking, and b) coordinating tasks to verbally fine-tune joint actions. Such a coordination task is the exit room mechanism (Figure 2, st. 2), for which two distant levers must be operated simultaneously (Figure 3, middle). The transcript shows alternating verbal suggestions and enquiries to coordinate actions.

L1: <i>Ok, wollen wir... Was steht da? Ziehen?</i>	L1: Ok, let's... What does it say? Pull?
L2: <i>Ziehen?</i>	L2: Pull?
L1: <i>Mhm.</i>	L1: Mhm.
L2: <i>Mhm. (pulls lever)</i>	L2: Mhm. (pulls lever)
L1: <i>Ich weiß nicht genau. (pulls lever)</i>	L1: I don't know exactly. (pulls lever)
L2: <i>Ja, funktioniert das bei dir?</i>	L2: Yes, does that work for you?
L1: <i>Äh, ich weiß nicht genau, ob ich richtig gezogen hab. Also es bewegt sich, bewegt es sich bei dir?</i>	L1: Uh, I'm not sure if I pulled it right. So, it moves, does it move at your end?

L2: *Nein. Oder sollen wir gleichzeitig das machen, oder?*
 L1: *Ja? Ich glaub schon.*
 [...]

 L2: *Eins, zwei, drei!* (both pull levers simultaneously)

 L2: No. Or should we do it at the same time, don't you think?
 L1: Yes? I think so.
 [...]

 L2: One, two, three! (both pull levers simultaneously)

Ultimately, simultaneous action execution is achieved by verbal synchronisation of the activity: L2: *Eins, zwei, drei!* (both pull levers simultaneously) [L2: One, two, three!]. – An example for a task that primarily elicits communication on collaborative problem understanding and solution construction, is the word puzzle (Figure 2, st. 4; Figure 3, left). The solution has to be placed with apples on lettered stone slabs.

L1: *Wollen wir erstmal gucken, was die Tafel sagt?*
 L2: *Sollen wir ein Wort buchstabieren, oder so?*
 L1: *Mhm! Mit dem Boden meinst du?*
 L2: *Es gibt, ja vielleicht, es gibt insgesamt fünf Äpfel. Stimmt? [...]*
 L1: *Ok. Wir müssen was buchstabieren.*
 L2: *Ja.*
 L1 (reads out the clue): *Der Apfel ist die Lösung.*
 L2: *Ja.*
 L1: *Wie verstehst du das?*
 L2: *Keine Ahnung* (laughs).
 L1: *Ok.*
 L2: *Aber, wenn es ein Lösung ist, so ein... ist das ganz direkt? Also der Apfel ist die Lösung?*
 L1: *Ja.*
 L2: *Also sollen wir APFEL buchstabieren, oder?*
 L1: *Ja, lass es ausprobieren.*
 L1: Shall we see what the board says first?
 L2: Should we spell a word or something?
 L1: Mhm! With the floor, you mean?
 L2: There are, yes maybe, there are five apples in total. Right? [...]
 L1: Ok We must spell something.
 L2: Yes.
 L1: (reads out the clue): The apple is the solution.
 L2: Yes.
 L1: How do you understand that?
 L2: I don't know (laughs).
 L1: Ok.
 L2: But, if it is a solution, such a... is that literal? So the apple is the solution?
 L1: Yes.
 L2: So we're supposed to spell APPLE, right?
 L1: Yes, let's try it out.

The design of the collaboration and coordination tasks seems to be responsible for eliciting the intended respective verbal behaviour.¹⁴

5.2.4 Affective Dimension: Fun

Finally, we want to examine the app's fun factor. The very fact that the tandems did not want to take a break within the first hour despite the recording directors' offers, shows that the learning was entertaining. In addition to frequent emotional expressions throughout the recording, especially a lot of laughter, we observed that particularly time-sensitive gaming tasks – like stealing apples and fleeing from the troll through the maze (st. 7, Figure 3, right) – elicited many affective expressions.

L1: *Welche ähm welche Richtung möchtest du gehen?*
 L2: *Ähm. Du kannst mal entscheiden.*
 L1: *Ok, dann gehen wir dahin, wo er grad nicht war.* (laughs) *Ehhh! Ogottogott. Ogottogottogottogott! Häh! Warum kann ich nicht weg? Warumkannnichtweg? Warumkannnichtweg?*
 L2: *Eeeecht?*
 L1: Which uhm which direction do you want to go?
 L2: Uhm. You can decide.
 L1: Ok, then we'll go where he wasn't just now. (laughs) *Uhhh! Dear God! (6x) Huh! Why can't I escape? (3x)*
 L2: For real?

In another example L1 expresses excitement by uttering fear, which, after both manage to escape, gets dissolved with L2's cry of victory.

L2: *Also, sollen wir gehen?*
 L1: *Warte, wenn er sich umdreht, oder?*
 L2: *Ok ja.*
 L1: *Also wenn er jetzt weggeht.*
 L2: *Jetzt?*
 L1: *Weiß ich nicht, also wenn er den Rücken zu mir. Ogottogottogott, ich hab Angst. Nein. oh o ä äh. Ok.* (manages to escape) *Bist du da?*
 L2 (also manages to escape): *Jahoiiii!*
 L2: Well, shall we go?
 L1: Wait when he turns around, right?
 L2: Ok yes.
 L1: So, if he walks away now.
 L2: Now?
 L1: I don't know, so when his back is directed towards me. Dear god! I'm scared. No. oh oh uh uh. Ok. (manages to escape) Are you there?
 L2 (also manages to escape): *Yahaaaay!*

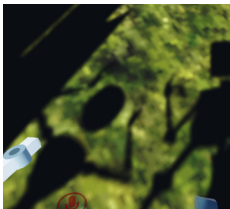
¹⁴ For epistemological reasons, the connection between cause and effect of course can never be shown.

5.2.5 Creative Embodied Interactions

In the break room (st. 6) in the middle of the labyrinth, the users explore further functions of the avatars and the environment. L1 draws L2's attention to the fact that the sun is visible in the sky, and both go to a sunny area and initiate a creative shadow play interaction. This could be interpreted as motor small talk to spend the mutual break time prosocially.

L2: *Guck mal, unser... Schatten?* [Look, our... shadow?]
(starts shadow play with her hands and arms)

L1: *Oh, stimmt! Oh mein Gott. Das ist voll komisch. Du hast einen runden Kopf.* [Oh, right! Oh my god. That's hilarious. You have a round head.] (joins shadow play)



L1: *Ich hab einen eckigen.* [I have a square one.]

L2: *Oh ja. Stimmt!* [Oh yeah. Right!]

L1: *Guck mal die Finger an, wenn die...* [Look at the fingers when they...] (shadow play with finger movements)

5.3 Discussion

This section will discuss the results regarding tandem learning with the *Hololingo!* app. Afterwards, we will classify the application with regard to Puenteadura's SAMR model on substitution, augmentation, modification and redefinition in language learning with social vr [22, 15].

5.3.1 Discussion of Results

The task design and narrative DGBLL realisation of the *Hololingo!* demonstrator showed the desired effects. Compared to self-directed hangout tandem data in *AltspaceVR*, the discourse participation of L2 was increased from 32 % to 46 %. The qualitative multimodal communication analysis showed rich communicative [19] and tandem-role adequate behaviour for L1 (adjusting pronunciation, empowering L2 to speak, providing guidance and feedback) and for L2 (taking the opportunity to speak, increasingly taking the lead, asking for feedback). The analysed coordination and collaboration tasks elicited the respective in-

tended verbal behaviour and therefore seem to be suitable for practising interactional skills. By engaging in gaming and problem solving they also relieved learners from topic construction. The analysis of affective expressions shows that the app provides a positive, entertaining atmosphere with excitement and fun in dynamic game sections.

Also the creative shadow play sequence in the stimulus-poor break room shows that even VR environments with moderate equipment and rough graphic implementation can offer opportunities for entertaining interactions, when it comes to Social VR: Despite the technological advances in immersion [23, 24], especially the interactive, playful multiplayer game character seems to facilitate the experience of presence (willing suspension of disbelief [12]) and of flow [21], as the tandems did not want to take a break during the first hour of use.

In another paper, we also evaluated user feedback [19]: L1 called the experience exceptional, enjoyable, and rich in variety. She also stated that working and solving tasks together helped getting to know each other. L2 claimed to have learned L1's way of thinking more directly through the setting and to have mentally anchored new words more deeply by speaking and acting simultaneously. However, L2 had expected more classical types of tasks such as cloze texts and formal assessments of her learning progress, so we may need to i) communicate the learning objectives more clearly, ii) support the habituation of the self-regulated autonomous method more strongly in the future, e. g. through a more detailed tutorial and improved progression tracking (note function, tests, badges). As the results reflect still an early stage of the app in the development process (demonstrator) and as they are based on a qualitative analysis of one DGBLL-tandem recording, they are limited regarding impact and replicability. More advanced versions of the app shall be evaluated with larger participant samples and complementary pre/post-tests.

5.3.2 Added Value of Digital Game-Based Tandem Learning in Social VR

In order to determine the added value of the digital method compared to analogue tandem learning, we adopt the four-stage model of Puenteadura [22, 15]. It assumes four levels of classification: a mere replacement without functional expansion (substitution), with functional expansion (augmentation), significant redesign (modification), and fundamentally new tasks (redefinition). Tandem learning in social VR with *Hololingo!* functions mainly as a direct substitute for analogue tandem learning on site. Although the

physical bodies of the learners do not meet virtually, the avatars linked to them by sensorimotor connections do. This makes joint leisure activities (hangout tandems) possible, but they may be experienced somewhat less personal or committed due to the lack of bodily closeness. At the augmentation level, the application can potentially connect a worldwide, simultaneously active pool of users that could not meet locally in analogue form. Since analogue tandems are self-directed, the game-based task design that facilitates language learning and communication can free the tandems from excessive autonomy and increase the awareness of the respective tandem roles. The task design can be used to target either holistic communication skills or specific language phenomena in connection with the respective curriculum. In the future, additional auxiliary and learning development measures, such as scaffolding and self-assessments, could transform analogue hangout-based language exchange into an efficient tandem learning experience. Also, by potentially being used as an assessment tool for teachers, some aspects of the application could be considered as redefinition of the tandem method.

6 Summary and Outlook

We presented the latest version of the *Hololingo!* demonstrator, a Social VR app for foreign language tandem learning. We i) argued for a didactic DGBLL approach to support the learning and training of verbal interactional communication competencies, ii) outlined the conception of different task types that we embedded in a narrative scenario, iii) described the DBR-based development and implementation process, and iv) carried out a linguistic evaluation: We examined a) the elicited discourse data for adequate tandem-role behaviour, b) effects of coordinative and collaborative task design, c) affective expressions as an indicator for excitement and fun and d) shadow play as a creative embodied interaction. To reduce dependencies on *VRChat*, we have implemented a new version in *Photon*. In addition to the demonstrated core design which augments, modifies and partially redefines tandem learning with regard to the SAMR model, we would like to explore further functions: We plan to implement scaffolding elements (notes, a dictionary, phrase suggestions, optional competence tests for individual progression assessment), create tasks to elicit specific language phenomena [4, 25], strengthen curricular connection to the CEFR, and add matching functions (user profiles, competence level indicator, badge system). In future studies, the linguistic

evaluation will be supplemented by a larger sample of participants for reasons of validity, by eye tracking to analyse (joint) attention processes and by pre/post-tests to measure learning effects and to study the relationship between game design and evoked speech.

Funding: Funded by the University of Hildesheim (starting grant) and the Stifterverband (program: Wirkung hoch 100, project: H110 5114 5132 36419, funding period: 12/2020–03/2022).

References

- [1] Ahlers, T., Bumann, C., Kölle, R., & Lazović, M. (2021). *Hololingo! – A Game-Based Social Virtual Reality Application for Foreign Language Tandem Learning*. In: Kienle, A., Harrer, A., Haake, J. M., & Lingnau, A. (Eds.): Die 19. Fachtagung Bildungstechnologien der Gesellschaft für Informatik, Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn, 37–48.
- [2] Ahlers, T., Lazović, M., Schweiger, K., & Senkbeil, K. (2020). Tandemlernen in Social-Virtual-Reality: Immersiv-spielebasierter DaF-Erwerb von mündlichen Sprachkompetenzen. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 25(2), pp. 237–269.
- [3] Ahlers, T. & Siegert, G. (2019). Sprachimmersion im Wohnzimmer: Interaktion, Grounding und Embodiment im DaF-Erwerb mittels Social-VR. In Philipp, H., Weber, B., & Wellner, J. (Eds.), *Kosovarisch-rumänische Begegnung. Beiträge zur deutschen Sprache in und aus Südosteuropa (FzDiMOS 8)* (pp. 94–117). Regensburg, Germany: Universität Regensburg.
- [4] Ahlers, T. (2018). Varietäten und ihr Kontakt enaktiv: Syntaktische Perzeptions- und Produktionsprozesse bei deutschsprachigen Zuwanderern in Österreich am Beispiel doppelter Relativsatzanschlüsse. PhD dissertation. Vienna, Austria: University of Vienna.
- [5] Anderson, L. & Krathwohl, D. (2014). A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of Educational Objectives. Edinburgh, UK: Pearson Education Limited.
- [6] Bechtel, M. (2010). Sprachentandems. In Weidemann, A., Straub, J., & Nothnagel, S. (Eds.), *Wie lehrt man interkulturelle Kompetenz. Theorien, Methoden und Praxis in der Hochschulausbildung. Ein Handbuch* (pp. 285–300). Bielefeld, Germany: Transcript-Verlag.
- [7] Biebighäuser, K. (2013). Fremdsprachenlernen in virtuellen Welten. Aufgabengestaltung in komplexen multimodalen Lernumgebungen. *Fremdsprachen lehren und lernen (FLUL)*, 42(2), pp. 55–70.
- [8] Bolton, S., Glaboniat, M., Lorenz, H., Perlmann-Balme, M., & Steiner, S. (2008). *Mündlich. Mündliche Produktion und Interaktion Deutsch. Illustration der Niveaustufen des Gemeinsamen Europäischen Referenzrahmens*. Berlin i. a., Germany: Langenscheidt.

- [9] Böcker, J., Ciekanski, M., Cravageot, M., Kleppin, K., & Lipp, K.-U. (Eds.) (2017). *Kompetenzentwicklung durch das Lernen im Tandem* (Arbeitstexte 29). Paris i. a., France: Deutsch-Französisches Jugendwerk.
- [10] Bumann, C., Kölle, R., Ahlers, T., Lazović, M., Elbeshausen, S., Schweiger, K., & Taranto, A. (2021). Das Geheimnis des Huckup – Ein Hildesheimer VR-Abenteuer zum DaF-Tandemlernen, unity-based vr demonstrator of an interactive communication game for tandem learning, updated version (available on request).
- [11] Council of Europe (2020): *Common European Framework of Reference for Languages*. Strasbourg, France: Council of Europe Publishing.
- [12] Dörner, R., & Steinicke, F. (2013). Wahrnehmungsaspekte von VR. In: Dörner, R., Broll, W., & Jung, B. (Eds.), *Virtual und Augmented Reality (VR/AR)* (pp. 33–63). Berlin, Germany: Springer Vieweg.
- [13] Ehlich, K. (2007). *Sprache und sprachliches Handeln. Bd. 1–3*. Berlin, Germany: De Gruyter.
- [14] Funk, H., Gerlach, M., & Spaniel-Weise, D. (Eds.) (2017). *Handbook for Foreign Language Learning in Online Tandems and Educational Settings* (Foreign Language Teaching in Europe 15). Frankfurt a. M., Germany: Peter Lang.
- [15] Hamilton, E. R., Rosenberg, J. M., & Akcaoglu, M. (2016). The Substitution Augmentation Modification Redefinition (SAMR) Model: a Critical Review and Suggestions for its Use. *TechTrends* 60. doi:10.1007/s11528-016-0091-y.
- [16] Hartfill, J., Gabel, J., Neves-Coelho, D., Vogel, D., Räthel, F., Tiede, S., Ariza, O., & Steinicke, F. (2020). Word saber: an effective and fun VR vocabulary learning game. In Preim, B., Nürnberger, A., & Hansen, C. (Eds.), *Tagungsband Mensch und Computer 2020*. (pp. 145–154). New York, USA: Association for Computing Machinery.
- [17] Heil, C. R., Wu, J. S., Lee, J. J., & Schmidt, T. (2016). A review of mobile language learning applications. *The Eurocall Review* 24(2), pp. 32–50.
- [18] Hung, H.-T., Yang, J. C., Hwang, G.-J., Chu, H.-C., & Wang, C.-C. (2018). A scoping review of research on digital game-based language learning. *Computers & Education* 126, pp. 89–104.
- [19] Lazović, M., & Ahlers, T. (submitted): DaF im Tandemlernen mit der *Hololingo!*-App. Eine Analyse von Tandemkommunikation in Game-based Social Virtual Reality.
- [20] McKenney, S., & Reeves, T. C. (2014). Educational Design Research. In Spector, J. M., Merrill, M. D., Elen, J., & Bishop, M. J. (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 131–140). New York, USA: Springer.
- [21] Nakamura, J., Csikszentmihályi, M. (2014). The Concept of Flow. In Csikszentmihályi, M. (Ed.), *Flow and the Foundations of Positive Psychology. The Collected Works of Mihaly Csikszentmihályi* (pp 239–263). Dordrecht, Netherlands: Springer.
- [22] Puenteadura, R. (2006). Transformation, technology, and education [blog post]. Retrieved from <http://hippasus.com/resources/tte/> (26.01.2022).
- [23] Slater, M., Lotto, B., Arnold, M. M., & Sanchez-Vives, M. V. (2009). How we experience immersive virtual environments: the concept of presence and its measurement. *Anuario de Psicología* 40(2), pp. 193–210.
- [24] Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology*, 109(3), p. 431–433.
- [25] Steels, L. (Ed.) (2012): *Experiments in Cultural Language Evolution* (Advances in Interaction Studies 3). Amsterdam/Phil., USA: Benjamins.
- [26] Zender, R., Knoth, A. H., Fischer, M. H., & Lucke, U. (2019). Potentials of Virtual Reality as an Instrument for Research and Education. *i-com* 18(1), pp. 3–15.

Bionotes



Timo Ahlers

Universität Potsdam, Potsdam, Germany
timo.ahlers@uni-potsdam.de

Dr. Timo Ahlers is a postdoc in the Department of Educational Sciences at the University of Potsdam. He studied General and Applied Linguistics as well as Cognitive Science at the University of Vienna, where he received his PhD in German Linguistics. His research interests include cognitive linguistics, language didactics, grammatical variation and educational technologies.



Cassandra Bumann

Universität Hildesheim, Hildesheim, Germany
bumann@uni-hildesheim.de

Cassandra Bumann studied Information Science at the University of Hildesheim. As a research assistant, she worked significantly on the development and implementation of the *Hololingo!* application.



Ralph Kölle

Universität Hildesheim, Hildesheim, Germany
ralph.koelle@uni-hildesheim.de

Dr. Ralph Kölle is a researcher at the Institute for Information Science and Language Technology at the University of Hildesheim. He studied computer science and completed his doctorate at the University of Hildesheim. His research focuses on virtual reality, e-learning and patent visualization.



Milica Lazović
Philipps-Universität Marburg, Marburg,
Germany
milica.lazovic@uni-marburg.de

Dr. Milica Lazović is a postdoc at the Institute for German Linguistics (AG DaF) at the Philipps-Universität Marburg. She studied German Linguistics and German as a Foreign Language and completed her doctorate at the University of Regensburg. Her research focuses on communication analysis, second and foreign language acquisition, advising in language learning, and research into language learning processes in digital learning environments.