Nordic Journal of Language Teaching and Learning

# (Dis)Engagement and Usefulness: Students' Experiences of Practicing Speaking German in a Virtual 360° Environment

**Elin Ericsson**
*University of Gothenburg*
elin.ericsson@ait.gu.se

**Sylvana Sofkova Hashemi**
*University of Gothenburg*
sylvana.sofkova.hashemi@gu.se

## Abstract

This study explored Swedish lower-secondary students' (*N*=41) experiences of practicing German speaking skills with filmed native speakers in a low-immersive virtual 360° environment. The students practiced dialogues in everyday-life scenarios, performing authentic tasks such as ordering in a restaurant. We employed various methods to produce data on students' self-reported experiences related to their speaking activity to gain a nuanced understanding of this way of practicing speaking. The data was gathered through logbooks, questionnaires, and interviews, which covered various aspects of the students' speaking activity. The twofold analysis consists of descriptive statistics on post-trial ratings, and thematic analysis based on the qualitative data. The findings show that the students mostly rated the overall experience as fun and easy. They estimated that they had developed their German speaking skills through practicing pronunciation and dialogue and felt better prepared for speaking. We show that some students felt disengaged even when participating in immersive simulations. There were individual differences regarding sustained engagement and experienced usefulness of the system, mostly linked to technical and contextual issues, being interrupted, misunderstood, feeling uneasy, or finding the speaking activity monotonous, too easy, or too hard. The study found that students sometimes completed dialogues without understanding the meaning of the words. We thus conclude with the need for pedagogical guidance and framing to embed this type of speaking activity in classroom practice.

**Keywords**: *conversational agent, foreign language learning, speaking skills, virtual 360° environment.*

 Nordic Journal of Language Teaching and Learning, Vol. 12, No. 1, 2024. ISSN: 2703-8629. DOI 10.46364/njltl.v12i1.1229

## 1. Introduction

Speaking a foreign language is considered challenging to teach and learn due to a lack of opportunities for spoken interaction in the target language during class (Håkansson Ramberg & Bøhn, 2023). Speaking is also acknowledged as causing emotional and cognitive barriers for students (Goh & Burns, 2012; Li, 2017). Still, the mastery of speaking in the target language coincides with how students measure, value, and align their success in language learning (Darancik, 2018). It is often the main aim of learning a language along with the desire to go abroad to where the target language is spoken is a highly rated motivational factor for studying languages (Dahl, 2022). Furthermore, being able to communicate in the target language is central in today's communicative and action-oriented approach in language education (Council of Europe, 2020). The importance of providing individual feedback and equivalent assessment for supporting this central aspect of language learning has been highlighted (Erickson, 2023).

Language education in Sweden is currently facing challenges with a high dropout rate of students in lower secondary schools and a lack of opportunities to practice spontaneous speech in authentic, everyday-life situations (Swedish School Inspectorate, 2022). In line with international research (Thiriau, 2017), Swedish language teachers have identified speaking as the most challenging skill to provide opportunities to practice in class, and the least supported skill by digital tools (Ericsson, Sofkova Hashemi, & Lundin, 2019). In their teaching, the Swedish language teachers face, for instance, big groups, lack of time, students' unwillingness to speak, and lack of natural contact with native speakers of the target language. We suggest this experienced lack of meaningful speaking practice with instant and individual feedback in the classroom and the students' inability to use the target language independently in everyday life might be changed with the introduction of emerging speech technologies for practicing speaking skills in a virtual learning environment (Meri-Yilan, 2019).

Virtual 360° environments, or computer-generated videos for physical presence in real or imagined environments for language learning, enable opportunities to interact orally face-to-face in simulated everyday-life scenarios with pre-recorded speakers of the target language (Berti, 2020). This video-based virtual reality (VR) technology promotes students to practice anytime and anywhere in simulated authentic places and situations (Golonka et al., 2014; Meri-Yilan, 2019; Mills, 2021). VR is a technology that offers simulated environments with varied and individualized learning opportunities in authentic learning situations that students can experience as real (Sun, 2017; Yang, et al., 2020). By replicating a genuine environment (Slater, 2018) through sound and visualizations, students become immersed (Parmaxi, 2023;

Wong & Notari, 2018) and feel safe to practice their target language speaking (Ayedoun et al., 2015; Thrasher, 2022). Research shows that practicing speaking in a VR environment boosts students' confidence and willingness to communicate in the target language (Papin, 2018). VR combined with automatic speech recognition, a technology that converts spoken language into written text allowing the students to interact with the pre-filmed native speaker through voice commands and dictation facilitates the teaching and learning of foreign language speaking skills, especially pronunciation (Bajorek, 2018; Golonka et al., 2014; Soto et al., 2020). Additionally, these technologies are beneficial for vocabulary development, listening skills (Soto et al., 2020), oral comprehensibility (Thrasher, 2022) and provide cultural awareness (Meri-Yilan, 2019) in meaningful (Mills, 2021) and "culturally authentic" speaking activities (Berti, 2020, p.327).

However, there is limited research focusing on how it is experienced by students in the language classroom (Parmaxi, 2023). Therefore, this study aims to explore self-reported student experiences related to the use of this kind of VR technology to introduce novel opportunities in an educational context for students to practice speaking skills in the target language. In the following sections, the areas of computer-assisted language learning (CALL) with regards to dialogue and VR will be explained, to contextualize the driving aims for this study, followed by a socio-cultural framing and an action-oriented approach.

## 2. Research on Dialogue-Based Computer- and VR-Assisted Language Learning

In the field of CALL, communicative and social approaches have gained recognition over form-based grammar teaching over the years (Thomas et al., 2012), promoting active student participation. Some widely accepted influential principles for effective second language learning concern "authentic input, conscious noticing on form, opportunity for interaction, in-time and individualized feedback, low affective filter, and an environment where language can be used" (Li, 2017, p.28). These established key principles highlight the importance of actively using the target language in meaningful, task-based activities (Ellis, et al., 2020) that center around topics relevant to students' daily lives and interests (Li, 2017), predominantly in interactional practices that prioritize meaning (Loewen & Sato, 2018). According to Bibauw et al.'s (2019) meta-analysis, interactional practice is enabled to various extents in spoken dialogue systems (SDSs) with conversational agents (CAs) that represent computer-controlled animated or pre-filmed artificial interlocutors. A CA interacts in such a system with the student as a "communicational counterpart" (Bibauw et al., 2019, p. 7) in the target language, always

accessible and never tired of repeating dialogues. Combined with VR, the interaction stimulates the students' abilities to listen, understand and respond to an additional language (Soto et al., 2020; Syderenko et al., 2018). CAs are represented in virtual contexts as simple text-based chatbots or visualized as animated virtual humans (Fryer et al., 2020). Various constraints on meaning and form regulate the spontaneity of the students' utterances in the dialogues (Bibauw et al., 2019; Fryer et al., 2020).

Adding VR to an SDS gives the opportunity to simulate physical presence, according to Yang et al. (2020), and the illusion of being on-site in the target language country (Mills, 2021; Lege & Bonner, 2020), which stimulates students' engagement (Dhimolea et al., 2022; Fryer et al., 2020; Wang et al., 2017), and enhances communication skills and cultural awareness (Berti, 2019; Lin & Lan, 2015). Such immersive potential of VR is recognized to provide "contextualized experiential language learning" (Dhimolea et al., 2022, p.1). For example, using the target language in a native-speaking environment to buy a cup of coffee in a café or buy a ticket in a museum. In high-immersive VR environments with the use of head-mounted displays, research shows that students take the VR simulation more seriously and spend more time on task (e.g., Alhalabi, 2016, Jensen & Konradsen, 2018). On the other hand, Makransky et al. (2019) found that although students are motivated and immersed in that kind of VR environment, they show poorer learning outcomes.

One example of low-immersive VR that enables students to overcome physical separation and feel immersed when learning (Berti, 2020; Meri-Yilan, 2019) is the desktop version of ImmerseMe (n.d.), a virtual 360° environment with filmed native speakers as CAs in photorealistic representations of everyday-life scenarios. This system's design aligns with active student-centered learning, encouraging students to take responsibility for their construction of learning through progression at their own pace (Meri-Yilan, 2019; Soto et al., 2020). Such variability could facilitate student spoken language learning on their own individual level, and important facet to explore in developing target language development practices (Håkansson Ramberg & Bøhn, 2023). As suggested by, for instance, Domínguez et al. (2013), some gaming elements are introduced into this non-game context (gamification), such as scores and levels, in order to increase students' engagement and promote the wished behaviors and enhance learning and development.

Recent research on VR-assisted language learning (VRALL) demonstrates that students experience practicing speaking in such systems as a low-stress and "fail-safe" learning environment (Soto et al., 2020, p. 9) indicating reduced affective filters in students (Lin & Lan, 2015), such as anxiety (Thrasher, 2022; Yang et al., 2020) and embarrassment, leading

towards greater confidence (Syderenko et al., 2018) and engagement to speak (Ayedoun et al., 2015; Papin, 2018). Experiencing engagement to speak is considered optimal and necessary for effective learning "characterized by enthusiasm, willingness, effortful exertion, interest and concentrated attention directed toward understanding, learning, or mastering the knowledge and skills necessary to be a competent language user" (Sang & Hiver, 2021, p. 32-33). VR environments are shown to promote such engagement by fostering motivation, interest, and excitement (Lege & Bonner, 2020). However, some recognized limitations of immersive technologies are, for example, pedagogical and individual diversity (Prensky, 2016), technical distraction (Berti, 2019), and low availability in schools due to high costs. Parmaxi (2023) highlighted in her review technical difficulties and systems being complex and time-consuming. As such, it is critical to better understand the experiences of students using these systems in class to determine the potential for utilizing them as supplementary language learning tools for interactive language *use* in the classroom in line with the communicative and action-oriented approach (Ellis et al., 2020; González-Lloret, 2017) as described also in European educational policy documents (e.g., Council of Europe, 2020).

### 2.1. Interaction in a Socio-Cultural Context and with an Active Approach

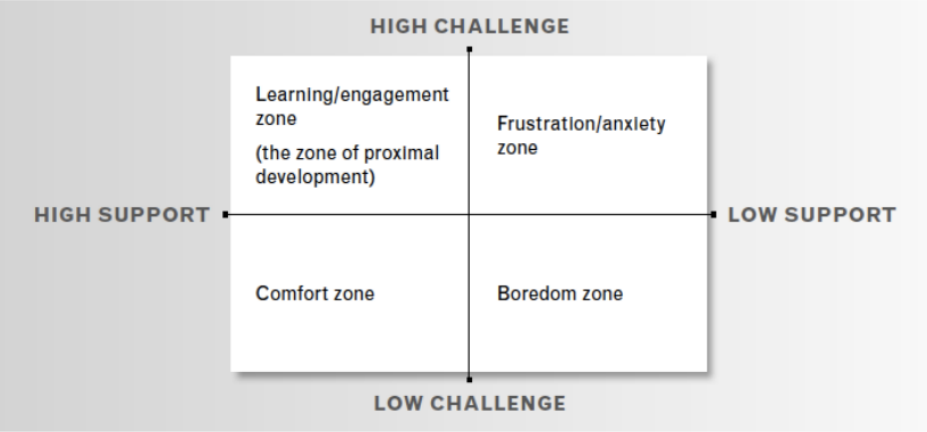
This study is based on the assumption that language learning occurs in a sociocultural context as active participation in interaction (Loewen & Sato, 2018) with more capable speakers, such as the teacher or other students (Lantolf et al., 2014). The filmed native speaker acting as CA in the simulated scenario represents the more knowledgeable and capable speaker. The SDS is offering temporary assistance that helps the student to complete the dialogue and move to new levels. This temporary assistance is referred to as "scaffolding" (Gibbons, 2015, p. 15), first introduced by Wood et al. (1976). The aim of this assistance is to challenge the student's cognitive limits to access the Zone of Proximal Development (ZPD; Vygotsky, 1978) to advance his or her potential level of competence, for instance, in speaking and cultural understanding (Lantolf et al., 2014).

It is recognized that students can achieve more when engaging in optimally challenging activities and within reach of assistance (Gibbons, 2015). Hence, we suggest using Gibbon's model adapted from Mariani (1997) with four zones of teaching and learning (learning/engagement/ZPD, frustration/anxiety, comfort zone, and boredom) as displayed in Figure 1, to interpret students' experiences in an SDS. This model relates

scaffolding to the level of the challenge of the task carried out. We have further adapted our version of the model in Section 5.

**Figure 1**

*Four zones of teaching and learning (Gibbons, 2015)*




*Note.* Gibbon's model (2015) was adapted from Mariani (1997).

Through accentuating communicative abilities beyond form-based approaches to language learning, DB-CALL promotes active and productive participation of students in authentic, interactive, and visually stimulating language learning environments (Bibauw et al., 2019; Thomas et al., 2012). Students who control their learning are in general active, which, together with motivation and self-confidence, are affective variables important language learning (Li, 2017). Additionally, in accordance with earlier VRALL research implicating VR's potential in SDS being beneficial for language learning, we introduced a virtual 360° environment.

### 2.2. Aim and Research Questions

Low-immersive VRALL studies focusing on speaking skills are still in their infancy (Lin & Lan, 2015). The majority are in higher education (e.g., Soto et al., 2020; Casañ-Pitarch & Gong, 2021; Meri-Yilan, 2019). To expand in this important area, this study explores how lower secondary school students in an institutional educational context experienced practicing German speaking skills with CAs in the desktop version of ImmerseMe (n.d.). In this study, we organize speaking skills into three subcategories: pronunciation, production of utterances, and interaction through verbal conversation. These categories are defined by the system's

ability to enable speaking practice with students and are used to organize our produced data on students' experiences. The research questions guiding this study are:

(i) How did the students experience practicing pronunciation, production of utterances, and constrained interaction in everyday-life scenarios with pre-filmed native speakers?
(ii) How did the students experience being challenged and scaffolded by the system in relation to their learning and development of speaking skills?

The overall aim of this study is to gain insights on how students' experience the use of CAs in SDS to offer suggestions for future use, and to better understand students' use of these tools in institutional educational contexts. The following section will outline how we produced data on these experiences, and how we organized the framing of the data.

## 3. Method

The current study was conducted in two classrooms with 13-to-14-year-old students who practiced speaking German as an additional language in a virtual 360° environment and instantly shared their experiences in logbooks, questionnaires, and interviews. This section presents the employed SDS and the design of our study, including the participating students, data instruments and analytical methods.

### 3.1. The Virtual 360° Environment

The speaking activities took place in the low-immersive desktop version of the virtual 360° environment in ImmerseMe (n.d.). It is an SDS with pre-filmed human native speakers of German acting as embodied CAs (Bibauw et al., 2019) in photorealistic representations of everyday-life scenarios that enable students to overcome physical separation and feel immersed when learning (Berti, 2020; Meri-Yilan, 2019). The system was selected based on the criteria of enabling immersed verbal face-to-face dialogues with CAs in foreign languages and performing tasks in everyday-life scenarios, as suggested in task-based teaching and learning (Ellis, et al., 2020). ImmerseMe provided lessons themed in 49 units about such topics as food, sports, and animals. The system applied Google automatic speech recognition in the dialogues, with explicit constraints on meaning and form, where the students repeated utterances from a

list of pre-set utterances in fixed paths (Bibauw et al., 2019). Spoken utterances had to be understood by the CA to keep the interaction alive and complete the dialogue. In line with constructivist ideas (Meri-Yilan, 2019), the student could choose from four learning modes and three levels (beginner, intermediate, and advanced) with a declining degree of scaffolding (from pronunciation to translation and immersion) to deconstruct the dialogue and assist the student, striving for full autonomy and ability to use the target language spontaneously.

The first mode (pronunciation) gave maximum scaffolding and displayed the exact utterance that the student was asked to read and reproduce aloud in drills. In the last mode (immersion), the student listened to the CA and, based on keywords, autonomously produced the memorized pre-set utterance in the target language. The speed of the CA's speech was modifiable. The system monitored the students how the system understood their pronounced utterances (text-on-speech), progress. As gaming elements create a sense of competition and increase engagement (Domínguez et al., 2013), the total score gained based on performance and accuracy of speaking was displayed. Additionally, there were reinforcements given through sound effects and words in bright colors flying around on the screen (e.g., KABOOM and WOW!). These elements of feedback and motivation factor into how students experience the speaking practice. The students' reported experiences will be described in the findings section.

### 3.2. Design of Study and Instruments

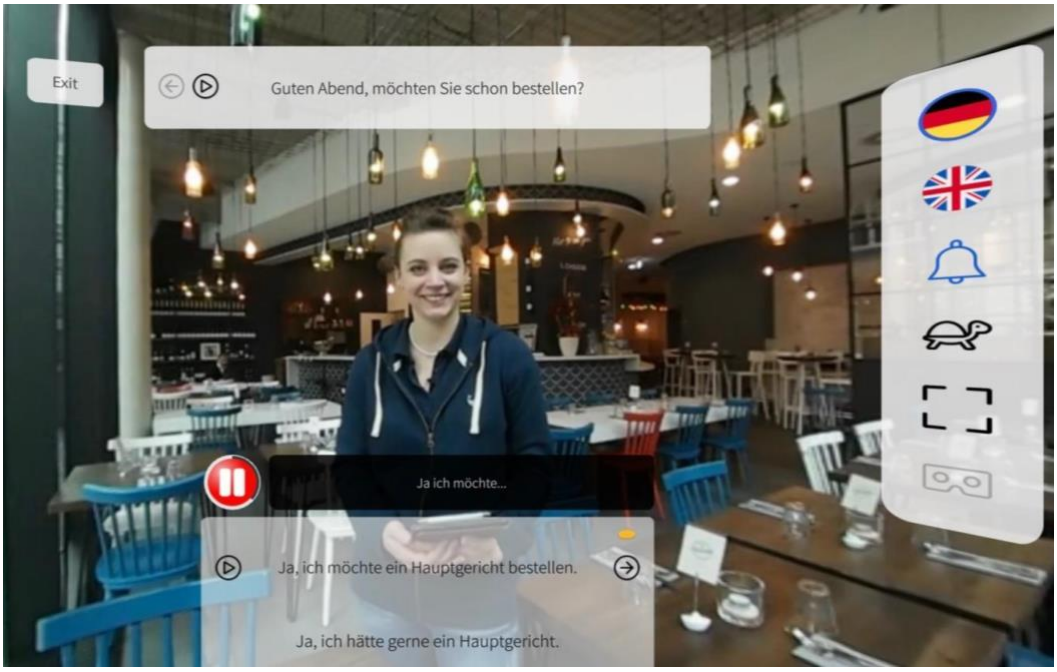
A convenience sample consisted of students aged 13 to 14 years *(N*=41). They were in two groups studying German at a beginner-elementary level (Council of Europe, 2020) for their second or third year as an additional language in seventh (*n*=25) or eighth (*n*=16) grade in a Swedish lower secondary school. All students had two hours per week at their disposal, divided over two lessons, dedicated to covering all skills including speaking, writing, listening, and reading. Before starting the trial, the students and their guardians signed informed consent forms as required by ethical guidelines for Internet research (Franzke et al., 2019). Overall, concerning the design of the study, we made decisions together with the teacher to adhere to the educational context with the existing lesson plan. The aim of the trial was to provide students with individual practice of speaking German by means of spoken interaction.

We provided the students with individual headphones with a microphone and logins to the system. The study entailed nine speaking sessions in the low-immersive desktop version of ImmerseMe (n.d.) over 18 days and the students experienced the 360° environment from

various angles by dragging the mouse pointer. The students practiced speaking German individually for approximately 15 minutes per session, performing everyday-life tasks while still maintaining a dialogue with the CA (Figure 2). The teacher guided the students with suggestions of scenarios, e.g., "Ordering a main," trying to match previously studied content of the course. The students also chose other scenarios at random according to individual preferences. Interacting with the system, the students were involved in speaking, listening, and reading to perform the given tasks, although the focus of this study was speaking.

**Figure 2**

*Screenshot of students' view: prompt to order food at a restaurant (ImmerseMe, n.d.)*




*Note*. © ImmerseMe Inc. Used with permission.

Initially, the students answered a web-based pre-trial questionnaire generating background information (see Table 1). After every speaking session, the students reflected upon writing down their immediate experiences in the SDS in shared documents functioning as digital logbooks. A final post-trial questionnaire was also provided, producing information on their experiences in the SDS, self-evaluation of speaking skills, and suggestions for improving the functionality of the system. A five-point Likert scale (Creswell & Creswell, 2018) was used

for rating, completed with open answers about what worked out well and poorly in the speaking practice with CAs and one multiple-choice item about what they estimated having practiced. The scale used the dimensions 1 = strongly disagree to 5 = strongly agree in reported statements. The questionnaire also asked about how it was to practice speaking in ImmerseMe (1 = very boring, to 5 = very fun), how it felt to practice (1 = very difficult, to 5 = very easy), how realistic the situations were (1 = not at all, to 5 = very much). The authors designed and pre-tested both questionnaires with inspiration in formulations from the Swedish research project about the connection between teaching, assessment, and learning (TAL) on speaking proficiency in language education (Erickson et al., 2022; TAL, n.d.) and an international snapshot study of student experiences of speaking practice with CAs in another SDS (Johnson, 2019). Finally, the first author conducted semi-structured interviews in Swedish with four volunteer students in pairs (post-trial) and their teacher (pre- and post-trial) until data saturation was reached to deepen our understanding of the learning situation based on the reported experiences. Representative parts of the interviews were transcribed. A selection of illustrative quotes from open answers in the post-trial questionnaire, logbook, and interview transcripts was conducted by the author. These quotes were translated from Swedish into English and presented along with the identified themes in Section 4.

**Table 1**

*Overview of data collection and analyses*

| | **Pre-trial Questionnaire** | **Post-trial Questionnaire** | **Logbook** | **Interview** |
|---|---|---|---|---|
| **Students (*N*)** | 41 | 39 | 36 | 4 |
| **Teacher (*N*)** | - | - | 1 | 1 |
| **Data** | Ratings<br>Open answers | Ratings<br>Open answers<br>Multiple-choice | Reflections<br>9<br>posts/person | Audio-recordings<br>Students (21min)<br>Teacher (70 min)<br>Transcripts |
| **Content** | Demography<br>Speaking<br>Digital tools | Experiences<br>Self-evaluation | Experiences | Experiences |
| **Analysis** | Descriptive | Thematic<br>Descriptive | Thematic | Thematic |

### 3.3. Data Production and Analysis

Empirical data in this study consists of students' self-reports in logbooks, ratings, free-text answers from the post-trial questionnaire, and interview transcripts (Table 1). This mix of data production methods provided an opportunity to tap into the students' experiences (Levy, 2015) from various angles to gain a deeper understanding of underlying aspects, a method commonly used in interaction-based research (Heift & Chapelle, 2012) within CALL (Bibauw et al., 2019).

Firstly, descriptive statistics give a quantitative overall understanding of students' experiences and self-evaluation. Secondly, reflexive thematic analysis was performed on collected data, combining a data-driven (bottom-up) approach with a theory-driven (top-down) approach (Braun & Clarke, 2019). The authors performed coding in Excel, applying colors representing what worked well, poor educational experiences, and improvement suggestions for this way of practicing in the SDS. Due to the complexity, to identify interrelated aspects, a visualization was created to facilitate the coding process. The authors compared and discussed until a collegial consensus was reached on the trustworthiness of the developed and reviewed themes (Nowell et al., 2017). Then a third researcher controlled and consented to the coding for inter-rater reliability. The thematic analysis yielded twelve themes (see Section 4.3) that converge into two overarching and interlinked categories: (Dis)Engagement and Usefulness.

## 4. Findings

This section first gives some background information about the participating students based on data produced in the pre-trial questionnaire before presenting findings based on the students' post-trial ratings of their experience of immersive spoken dialogues and what they estimated having practiced and developed. Next, we proceed with the findings of the reflexive thematic analysis, based on data derived from open answers in the post-trial questionnaire and logbook. There is an initial visualization of the complexity of identified aspects of experiences organized under the categories '(Dis)Engagement', represented by seven themes, and 'Usefulness', which yielded five themes, presented without inner ranking order. Selected quotes from open answers in the post-trial questionnaire, logbook reflections, and transcribed interviews are presented along the themes.

### 4.1. Background Data about the Students

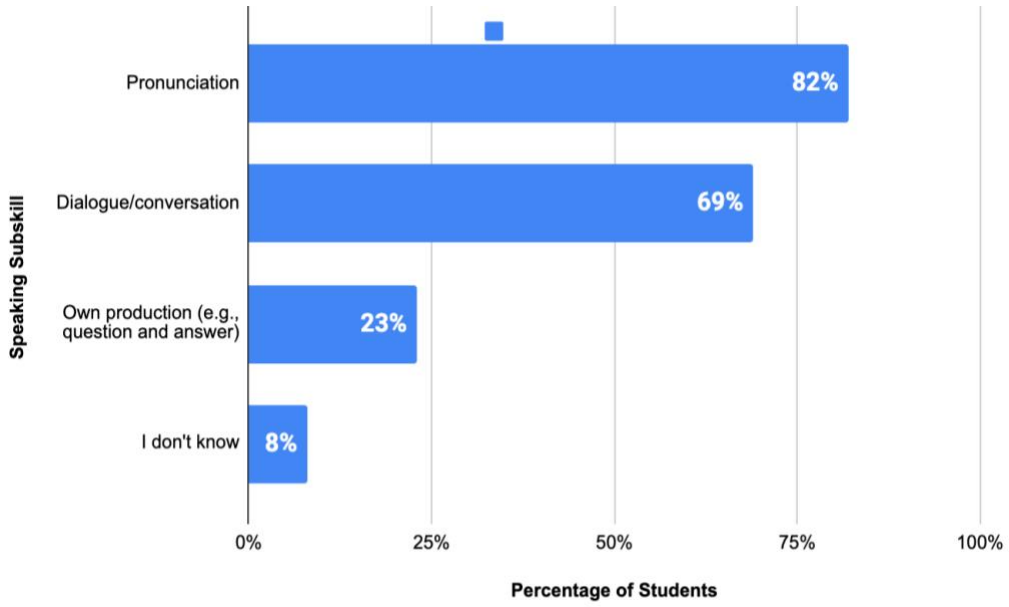
In the pre-trial questionnaire, the students reported spending approximately seven hours daily on digital devices. The majority (76%) were positive about using digital tools that aimed to enhance the process of learning German, such as speaking (59%) with CAs in VR (59%). Few students reported learning German as boring (19%) and difficult (19%). Even fewer (14%) reported speaking it outside school.

### 4.2. Post-Trial Evaluation Ratings

Results showed that most students in the post-trial questionnaire rated their summarized overall experience positively as good/very good (64%) in relation to the development of speaking skills, fun/very fun (59%), and easy/very easy (62%), and with a low threshold to get started with the system (63%). Few voices complained of a bad (5%) or boring/very boring overall experience (13%). Contrastingly, 44% reported the speaking scenarios as highly realistic/very highly realistic (like in everyday life). The students' self-estimation of the practiced subskills revealed mainly "pronunciation" (82%) and interaction in "dialogue and conversation" (69%) (see Figure 3). Almost a fourth of the students reported "own production" in such questions and answers (23%). A minority of the students (8%) did not know what they had practiced.

**Figure 3**

*Students' self-estimation of practiced subskills*



Additionally, the post-trial self-evaluative results demonstrate that the majority of the students experienced speaking German in an SDS as a beneficial learning activity, as presented in Table 2.

**Table 2**

*Students' post-trial self-evaluation*

*[After these weeks' trial of practicing speaking German with ImmerseMe…]*

| **Item** | **Statement** | ***Mdn*** | ***Min*** | ***Max*** | ***%*** |
|---|---|---|---|---|---|
| 1. | I have improved my German speaking skills. | 4 | 1 | 5 | 59 |
| 2. | I feel less anxious about speaking German. | 4 | 2 | 5 | 51 |
| 3 | I am more motivated to speak German now than pre-trial. | 3 | 1 | 5 | 38 |
| 4. | I am more prepared for everyday-life speaking situations now than pre-trial. | 4 | 1 | 5 | 54 |
| 5. | It has been good for me to speak with "native speakers." | 4 | 1 | 5 | 56 |
| 6. | It has been good for me to speak in different everyday-life situations. | 4 | 1 | 5 | 54 |
| 7. | It has been good for me to practice listening and understanding German. | 4 | 2 | 5 | 64 |

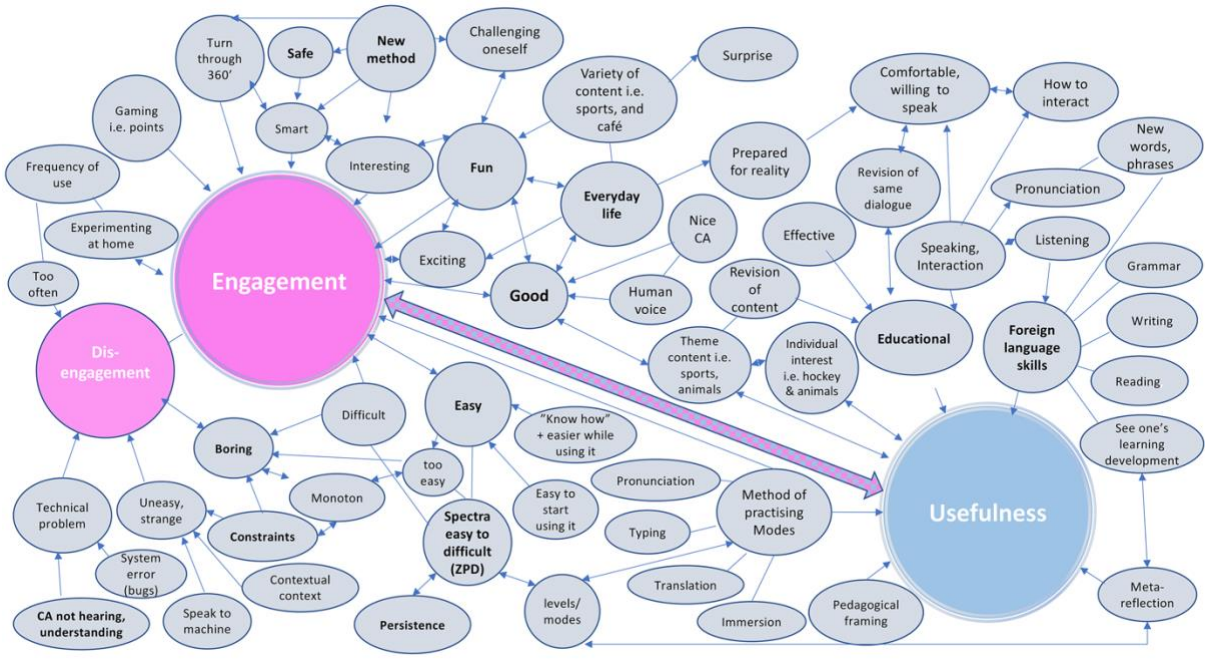
*Note*. The five-graded Likert-scale ratings went from 1 = not at all, and 5 = to a very great extent. Percentages represent students whose ratings were positive (4 or 5).

### 4.3. Reflexive Thematic Analysis – Complex Aspects

Based on data derived from open answers in the post-trial questionnaire and logbook, Figure 4 represents a visualization of the emergent coding process of the thematic analysis. The figure reveals the complex spectra of interrelated aspects of students' self-reported experiences that were identified and organized into two categories: (Dis)Engagement and Usefulness, all concerning their learning and development of speaking skills.

**Figure 4**

*Coding spectra of aspects of self-reports in a spoken dialogue system*




*Note*. This figure does not claim to be proportionally designed. Aspects marked in bold are interpreted as central to the students' experiences in the virtual 360° environment.

### *4.3.1. (Dis)Engagement*

In this sub-section, we present the seven themes addressing various interrelated aspects of students' experiences associated with the spectra of **engagement** and its opposite, **disengagement.**

**Degree of enjoyment and realism** represents the first theme with students reporting the ways of practicing as "cool," "special," "exciting," "fun," "interesting," and "lively," especially concerning the promotion of "realistic," "authentic," "real life," and "real-time" scenarios interacting with a person:

> It did not become so boring; (…) there was a person you spoke with and a screen where you could see, like being there in real life, so it was not just like speaking with a computer and not having a clue what to do later when you actually are in a café. They had really succeeded in making it realistic indeed. (...) You can walk around, and it speaks in real-time. (Student 3)

The students highlighted that they felt immersed, stressing how movement in the environment makes it livelier:

> It becomes more realistic when you sit with a person and look around in the room. It feels like you are within that world when speaking with that person. However, some [CAs] do not use the environment; they just sit in front of you, but some people walk around, which is cool. (Student 25)

Also, the teacher summarized that the multidimensional environment would preferably be used more actively, both culturally and interactively, by providing "something clickable, like a menu on the wall, or small natural dialogues with other people around you within the scenarios (...), in order to keep up the students' interest" and motivate them to move around.

Some students voiced signs of being less enjoyed: "It was not so realistic either, and the interaction did not feel good" (Student 30). They experienced the sessions as simple and tedious, wishing for more fun elements: "It is a very monotonous teaching material since there is only one learning thing. Add more game modes, and one will find it more fun for a longer time" (Student 28). The gaming elements seemed to stimulate many students who regularly reported how many points they had gained, although the teacher and some students expressed concerns about the scoring model. Some students reported using the system outside the educational context; "Today I achieved 93 points, but it is because I have practiced at home" (Student 25), where they also tried other languages such as English, Chinese, and Spanish.

Secondly, the students experienced the scenarios as having **varied content** in comparison to earlier study material in class, also linking to the students' interests: "I have done sports today, it was fun since I love sports, but it would have been better if you could talk about hockey" (Student 7). Besides sports, students also commented frequently on food and animals, and the wide range of themes was appreciated. However, the teacher pointed out some inappropriate scenarios for the age group, such as buying beer. Additionally, practicing directions when being chased in a street in Germany, was considered to give the wrong cultural association, and a suggestion would be to "ask for the way where to buy a wurst [sausage] instead to include the German food culture" (Teacher).

In the third theme about **impressions of CAs,** the students highlighted the CA's as being "nice" with a "humanlike voice," "good at understanding," "guessing words," "comprehensible," and "moving around." On the other hand, some students also found the CAs "strange," "frustrating," and that they felt "uneasy" speaking to them. One student

expressed that "they speak damn slowly, so it takes an eternity" (Student 38) and another that "if you failed one word then the computer failed … no, no then you will have to restart again" (Student 3). The CAs' sensitivity to surrounding noise was experienced differently: "It understood everything I said even when the person next to me was talking, which was impressive" (Student 40) and "the computer heard correctly more often [today] because I did not sit close to anyone (Student 12). Some students found the requirement to repeat preset answers rather than produce their own utterances spontaneously very repetitive and boring: "I would like to speak with my own words to the computer, so it understands what I say, real communication" (Student 2).

The fourth theme, **degree of comfort and safety** concerning surroundings in the classroom highlights that some students experienced this as making them feel "uncomfortable," especially initially: "I do not like doing this in the classroom since it feels that others listen to me when I am speaking. It is uncomfortable" (Student 20). However, many students expressed that it felt "safe" speaking with the CAs. They felt more comfortable speaking using the system compared with traditional classroom speaking activities:

> Maybe I have never spoken to that person, so maybe I do not feel like having a conversation for five minutes about what I am going to have for dinner (in German), so of course, you feel more comfortable with a computer. (Student 3)

**Being unable to finish dialogues and tasks properly** is the fifth theme, which is mainly about the inability of the CAs to hear or understand. The students rarely suspect it is *their* pronunciation that is at fault: "It went so-so considering that the idiot did not understand what I said" (Student 3). In fact, many of them blamed the CAs, the system, the contextual noise, or the microphone. The students felt "frustrated," "annoyed," "irritated," or "bored," when even though they were expressing the utterances correctly, the CAs still interpreted it wrongly: "I worked with Backpackers, and it went well and was fun, but it did not work properly when I said *Dir auch*, so it was not there, or it did not enter the tool, so we reported an error" (Student 1). Many students also experienced being interrupted too early: "The robot/AI interrupted me before I had time to finish speaking, which resulted in fewer pronunciation points" (Student 11). Frequently commented upon are technical issues with the system such as "bugs," "lagging system" with unstable functionality such as "buttons," "scrolling down," the CA saying something wrong or interpreting something wrongly, and hence the "inability to finish tasks

properly." The teacher's concluding remark was: "It is enough that it malfunctions once; the students condemn the whole tool [...] there are no grey areas, just 'excellent' or 'crap'."

**Degree of challenge** represents the sixth theme with diverse spectra of difficulty levels leading to (dis)engagement. Many students reported that it was "easy" to practice in the system but also: "Some exercises were quite boring since they were somewhat too easy" (Student 14). This result could be because they were already familiar with the content: " Since we have already worked pretty much with that [what I like eating], it was somewhat too easy" (Student 16). Many students experienced being challenged at different levels: "Clothes were on beginner-level, but the exercise about tourist information was on an intermediate level. It was fun to be challenged, and you could talk a lot, which is good" (Student 27).

Students also felt challenged by the different modes: "I worked with Museum, and it was not very difficult, but it was more difficult with translation and immersion" (Student 4). Here is a student reflecting on scaffolding and progression to an adequate level in relation to developing speaking skills:

> When you sit with your peers, you need to stop and think about how it is pronounced, but here you have everything in front of you. You know how it is pronounced, which gives you better fluency, which also makes you more comfortable, and that you dare more and then advance to more difficult levels where you translate and have keywords. (Student 3)

**Time and frequency of practice** are highlighted aspects that influenced the students' experiences: "It is a good idea, but it is all about how you use time and how often you should work with it" (Student 36). With enough time spent getting involved in the scenarios, the students initially expressed engagement. On the other hand, linked to disengagement, they experienced the system as being more "monotonous" and "boring" when used "too frequently" and towards the end of the trial, when the system lost its novelty effect: "It was fun in the beginning when it was a new way of working. (…) It quickly becomes boring, though, and after a few lessons you are tired of it" (Student 28).

### *4.3.2. Usefulness for the Development of Language Skills*

In this part, we address interrelated aspects experienced by the students as beneficial for learning and developing foreign language skills (especially speaking), presented within five themes.

The students believed they **improved their speaking skills and gained confidence** through everyday-life scenario practice using their speaking, listening, reading, and writing skills. The students frequently reported the activity as "effective," "educational," "powerful," "smart," and "useful." The students recognized improving and enjoyed practicing pronunciation: "I think it was a fun and easy way to become better at speaking German and learning how you pronounced things you earlier have pronounced wrongly" (Student 21).
They talked about deployed strategies: "You must speak clearly and not too fast, so the computer follows. Yes, you think about pronunciation" (Student 3). Students also mentioned the scaffolding provided: "You could listen to what you should say if you did not know how it was supposed to be pronounced" (Student 8). "Having conversations in the target language felt satisfying: It helps you to be more comfortable with speaking German (and other languages)" (Student 29). This also prompted reports about "better fluency" and "gaining confidence":

> I have spoken much more. Otherwise, it is not very often you practice speaking, you are at different places [in teaching material], so consequently, you skip speaking exercises (...). Now I have more fluency and (...) it has really helped me practice speaking more German. (Student 3)

Many students reported feeling better prepared for everyday-life conversation: "I am better at speaking with German people in a real-life scenario, and I can understand somewhat more German" (Student 11). "I have learned courtesy phrases which can facilitate a lot if you go to, e.g., Germany one day" (Student 2).

The second theme, **learning new words without always understanding their meaning**, raised concerns among the students: "I think I learn pronunciation, but mostly I do not know what I am saying" (Student 16). However, they still manage to complete the dialogues. The students were creative in suggesting solutions, such as "a translation after I have answered" (Student 5) or working interactively with "clickable vocabulary visualized in the environment," for instance, the menu.

The third theme concerns the students experiencing the system as their **personal teacher with instant feedback**. Instead of traditionally one teacher present in the class, the students experienced a "new dimension" and "different from traditional teaching material" of "thinking in a new way" about learning when having one native speaker to guide them: "You have your own teacher there who can say exactly how it is pronounced, so of course it helps" (Student 3). They experienced instant feedback: "You saw immediately what you had said wrong since it was marked in red". (Student 27). However, the class teacher expressed concerns about the students "getting away with making many mistakes without being corrected," giving examples such as significant phonemes (/z/ versus /ts/). Hence, the teacher suggested follow-up exercises focusing on pronunciation by "drilling the correct version," also echoed by some students' wish for more corrective feedback on pronunciation to highlight their gaps in knowledge. Additionally, the visualization of fulfilment of dialogues and gained score in the system was experienced useful in this learning situation: "It was actually somewhat fun to see how much you had developed during the last weeks" (Student 20). The students expressed their wish for further development of that feature.

Students reported being **non-distracted, independent, and self-regulating** in their learning. They found themselves spending more time speaking German and "concentrating," cut off from the others in the classroom since they had headphones on. Hence, they were not distracted by interactions in Swedish with their classmates. Moreover, students highlighted it as beneficial to be able to regulate the activity themselves: "there are several levels and different modes to use with the different lessons, which I like" (Student 29).

Finally, the fifth theme includes **revision and repetition.** The students made connections to previously gained knowledge, exemplified by revision of familiar content and vocabulary: "The tool is also very useful for revising something that you have been taught during the lesson" (Student 31). Also, unlimited repetition of the same dialogue for perfection was reported as beneficial: "Working with The Family, I was somewhat surprised since I had a lot to revise although I knew everything beforehand" (Student 20).

## 5. Discussion

The study explored how lower-secondary students in an institutional educational context experienced practicing German speaking skills in the virtual 360° environment ImmerseMe. In terms of answering the research questions, the first section is about the practice of speaking in

the system and the second section discusses the results in relation to experienced scaffolding and challenge.

### 5.1. (Dis)Engaging and Useful

The study demonstrates that overall, the participating adolescent students experienced practicing speaking skills in dialogues with the CAs as good, fun, easy, realistic, and immersive, pointing towards engagement, in line with results in higher education studies (e.g., Casañ-Pitarch & Gong, 2021; Meri-Yilan, 2019). This may also be explained by students' initial positive views (Levy, 2015), here of learning German and of using digital tools for learning. However, the experience of using the system during the whole period varied from continuous engagement to gradual disengagement, manifested as boredom or frustration for a while, often linked to technical issues with the system or the CAs not understanding them properly.

The students further highlight that they felt better prepared for real-life conversations, gaining confidence, and experiencing less anxiety about speaking, which is in line with earlier research on spoken interaction in VR (Ayedoun et al., 2015; Thrasher, 2022), also including low-immersive VR (Papin, 2018). The study shows that their degree of engagement was related to the experiences of being immersed (Meri-Yilan, 2019; Wang et al., 2017) and being able to move around, and, for some, being further motivated by the gamification (Domínguez et al., 2013). Engagement was also linked to experiencing the content as varied and relating to the students' interests or topics with which they were already familiar with (Li, 2017). Practicing in a safe environment (Soto et al., 2020) made the students feel primarily comfortable, with humanlike, nice CAs, except when the CAs were not able to hear, understand, or interpret the students (Fryer et al., 2020).

The students reported practicing speaking skills, mostly pronunciation and interaction in dialogues. On the other hand, production of own utterances was low and experienced as repetitive and rather drill-like than constructive (Berti, 2020) due to the system's constraints with pre-set utterances (Bibauw et al., 2019). The speaking sessions were therefore mostly found to be useful and comfortable to practice pronunciation (Bajorek, 2018; Meri-Yilan, 2019) and vocabulary (Soto et al., 2020). However, students also reported using new words without always understanding their meaning. In contrast to the complex issue of feedback and assessment in traditional teaching of speaking (Erickson et al., 2022; Erickson, 2023), these findings point toward the usefulness of instant feedback and assessment displayed by the

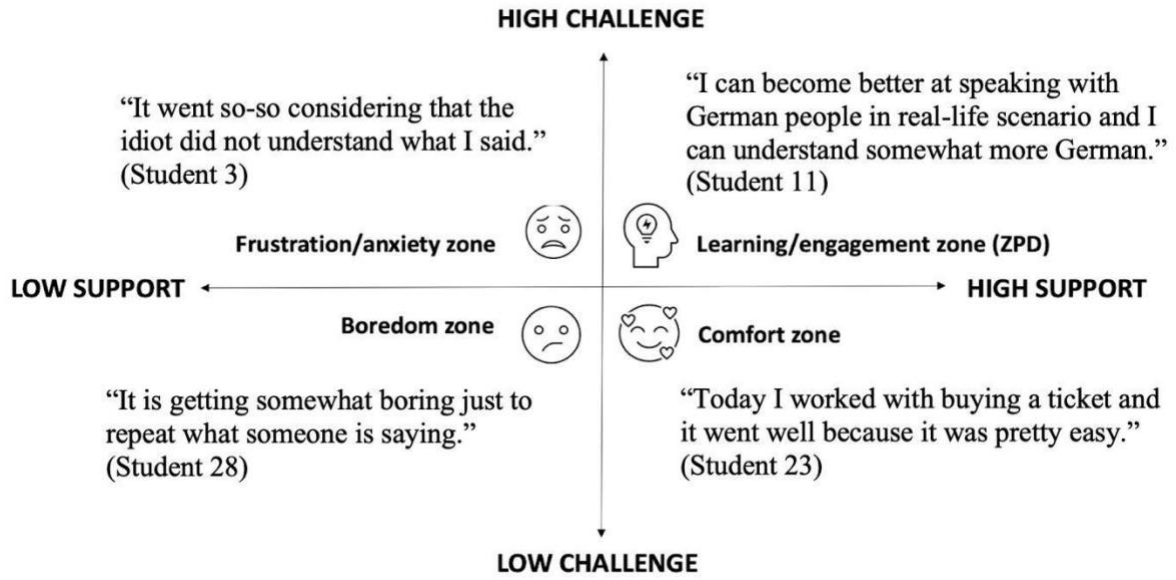
system through such as transcripts and gained scores based on performance and accuracy, which could facilitate for both teacher and students.

### 5.2. Challenged and Scaffolded in Four Zones

We further interpret the students' experiences in accordance with the kind of scaffolding and challenges they experienced, applying Mariani's (1997) and Gibbons' (2015) four zones of teaching and learning (Figure 5). Occasionally some students experienced being too challenged in combination with a level or mode too advanced, giving them too little support from the system and resulting in going back one level, changing the progression mode in the system, or giving up. In contrast, other students aware of their learning development (in the ZPD) found that higher challenges helped them to practice and develop their speaking skills and to engage fully with their hearts, and minds (Sang & Hiver, 2021). Some students reported boredom when engaged in exercises lacking enough support or challenge and experienced the reproduction of pre-set utterances as monotonous. Therefore, if they did not find the speaking activity challenging, there was consequently not so much learning going on (Gibbons, 2015; Lantolf et al., 2014). The students differed in persistence when meeting various difficulties. Some lost interest immediately, condemning the tool whereas others continued their practice. Hence the teacher could be of help to guide the students further.

**Figure 5**

*Visualization of provided scaffolding in relation to experienced challenge*



*Note.* Our adaption is based on a model from Mariani (1997) and Gibbons (2015), here visualizing an interval from low support in the SDS on the left-hand side to high support on the right-hand side.

The students felt able to pace their own learning process (Meri-Yilan, 2019) while receiving feedback from the system and the CAs, though some felt frustrated by the constraints of the system and the drill-like focus of the practice. The students also expressed frustration when the CAs did not hear or understand and misinterpreted them (c.f. Fryer et al., 2020). It was, however, rare for the students to suspect that it was their own part in the dialogue that caused communication breakdowns, with most blaming the system or the CAs. Since the system does not enable spontaneous speech, students are not able to improvise their utterances, although they may know how to pronounce them accurately. This might be a challenge for high-proficiency students who are used to using more elaborate language. However, if they add extra words, then they are not understood by the system. Additionally, since the students do not receive detailed feedback on the kind of mistakes they make in their spoken utterances, neither on grammar nor on their pronunciation, the students do not comprehend their eventual part in the breakdown, what they need to develop, or if it was a real bug or technical issue of the system.

## 6. Conclusion

This study indicates that practicing speaking skills in a low-immersive virtual 360° environment with CAs has a promising potential (c.f., Bajorek, 2019; Berti, 2019; Papin, 2018) in terms of making this age group gain confidence and feel better prepared for everyday-life conversation in a foreign language. Results of the study reveal that students' experiences varied from continuous engagement in fun and easy speaking activities to gradual disengagement experiencing boredom and monotonous practice. This indicates various levels of persistence to keep engaging in the spoken interaction with the CA, tightly linked to the degree of challenge and scaffolding that coincides with the individual student's ZPD. Additionally, the students experienced receiving useful feedback and assessment on their performance, above all when it concerns pronunciation, which could add to the call for feedback and equivalent assessment on speaking a foreign language (Erickson, 2023).

Limitations of the study relate above all to the current educational context. It is unclear to what extent our findings generalize to other contexts with other possibilities of natural

contact with the target language, other proficiency level and ages of the participating students, content orientation and integration of different language skills. Further studies are needed to determine whether, in fact, communication breakdowns are due to students' non-adequate level of speaking proficiency, the system's constraints, failing speech recognition or unexpected technological issues. Additionally, since VR technology is evolving rapidly (Lege & Bonner, 2020) there is a need to explore students' experiences, also in high-immersive VR (Dhimolea et al., 2022), and the relevance of authentic cultural situations for learning. We suggest longitudinal studies to exclude novelty effects (Fryer et al., 2020) and explore correlations in triangulated data around experiences, performance, and time on task.

To conclude, this study demonstrates that this age group of students feel immersed when practicing pronunciation and spoken interaction in the virtual 360° environment. It indicates the need to make even better use of simulated immersive contexts involving more cultural aspects (Millis, 2021; Berti, 2019) and interactivity for vocabulary practice. It further implicates the importance of individual and pedagogical framing (Heift & Chapelle, 2012; Lege & Bonnier, 2020; Parmaxi, 2023) when embedding SDS with VR in classroom practice. To avoid unnecessary frustration of the students, these findings could help teachers to gain a pre-understanding of the aspects of importance of an SDS and how to guide students into the right balance of challenge and support for beneficial engagement. Practical implications of this study are also that the students' educational experiences are overall in alignment with established key principles for adequate communicative foreign language teaching and learning (Ellis, et al., 2020; Li, 2017), which indicates that this is a beneficial method for practicing speaking. Additionally, the results inform instructional developers on how to optimize these systems with fewer constraints, age-adapted content, and more elaborated individual feedback on pronunciation and development, enabling students at an early age to practice foreign language speaking skills to be prepared for spoken interaction in real life.

## 7. Acknowledgements

This study has been made possible thanks to ImmerseMe, so major acknowledgment is due to Scott Cardwell, co-founder, and CEO. Many thanks, not least to the participating students and their teacher.

## 8. References


Alhalabi, W. S. (2016). Virtual reality systems enhance students' achievements in engineering education. *Behaviour & Information Technology, 35*(11), 919–925. https://doi.org/10.1080/0144929X.2016.1212931

Ayedoun, E., Hayashi, Y., & Seta, K. (2015). A conversational agent to encourage willingness to communicate in the context of English as a foreign language. *Procedia Computer Science*. *60*(1), 1433–1442. https://doi.org/10.1016/j.procs.2015.08.21

Bajorek, J. P. (2018). *L2 Virtual reality-based ImmerseMe software: Speaking, interactivity, & time [oral presentation]*. Computer Assisted Language Instruction Consortium (CALICO).

Berti, M. (2019). Italian open education: virtual reality immersions for the language classroom. In A. Comas-Quinn, A. Beaven & B. Sawhill (Eds.), *New case studies of openness in and beyond the language classroom* (pp. 37–47). Research-publishing.net. https://doi.org/10.14705/rpnet.2019.37.965

Berti, M. (2020). Learning technology review: ImmerseMe. *Calico Journal 37*(3), 321–330. https://doi.org/10.1558/cj.39714

Bibauw, S., François, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning*, *32*(8), 827–877. https://doi.org/10.1080/09588221.2018.1535508

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise, and Health*, *11*(4), 589–597. http://doi.org/10.1080/2159676X.2019.1628806

Casañ-Pitarch, R., & Gong, J. (2021). Testing ImmerseMe with Chinese students: acquisition of foreign language forms and vocabulary in Spanish. *Language Learning in Higher Education*, *11*(1), 219–233. https://doi.org/10.1515/cercles-2021-2016

Council of Europe (2020), *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*, Council of Europe Publishing.

Creswell, W. J., & Creswell, D. J. (2018). *Research Design, Qualitative, Quantitative and Mixed Methods Approaches (5th ed.).* Sage.

Dahl, A. (2022). University language students' motivations for their language of study. *Nordic Journal of Language Teaching and Learning*, *10*(1), 1–23. https://doi.org/10.46364/njltl.v10i1.1013

Darancik, Y. (2018). Students' Views on Language Skills in Foreign Language Teaching. *International Education Studies*. *11*(7), 166–178. https://doi.org/10.5539/ies.v11n7p166

Dhimolea, T. K., Kaplan-Rakowski, R. & Lin, L. (2022). A Systematic Review of Research on High-Immersion Virtual Reality for Language Learning. *TechTrends, 66*(5), 810–824. https://doi.org/10.1007/s11528-022-00717-w

Domínguez, A., Saenz-de-Navarrete, J., De-Marcos, L., Fernandez-Sanz, L., Pages, C., & Martinez-Herraiz, J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers and Education, 63*, 380–392.

Ellis, R., Skehan, P., Li, A., Shintani, N., & Lambert, C. (2020). *Task- Based Language Teaching: Theory and Practice.* Cambridge University Press.

Erickson, G., Bardel, C. Österberg, R. & Rosén., M. (2022). Attitudes and ambiguities – teachers' views on second foreign language education in Swedish compulsory school. In C. Bardel, C. Hedman, K. Rejman & E. Zetterholm (Eds.), *Exploring*

*Language education: Global and local perspectives* (pp. 157–223). Stockholm University Press.

Erickson, G. (2023). Bedömning och återkoppling – med fokus på muntlighet. *Nordic Journal of Language Teaching and Learning (NJLTL), 11*(1), 70–88. https:/doi.org/10.46364/njltl.v11i1.1037

Ericsson, E. Sofkova Hashemi, S. & Lundin, J. (2019). Teachers' attitudes and experiences of digital tools within speaking proficiency in foreign teaching and learning – Interactional competence left in the dark. In *Proceedings of ICERI 2019, 12th annual International Conference of Education, Research and Innovation* (pp. 861-870). IATED. https://doi.org/10.21125/iceri.2019.0272

Franzke, A. S., Bechmann, A., Zimmer, M., & Ess, C. (2019). *Internet research: Ethical guidelines 3.0: Association of internet researchers*. https://aoir.org/reports/ethics3.pdf.

Fryer, L. K., Coniam, D., Carpenter, R., & Lăpușneanu, D. (2020). Bots for language learning now: Current and future directions. *Language Learning & Technology, 24*(2), 8-22. http://hdl.handle.net/10125/44719

Gibbons, P. (2015). *Teaching English language learners in the mainstream classroom*. Heinemann.

Goh, C. C. M., & Burns, A. (2012). *Teaching speaking: A holistic approach*. Cambridge University Press.

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L. & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning*, *27*(1), 70–105. https://doi.org/10.1080/09588221.2012.700315

González-Lloret, M. (2017). Technology for task-based language teaching. In C. A. Chapelle, C. & S. Sauro (Eds.), *The handbook of technology and second language teaching and learning* (pp. 234–247). Wiley-Blackwell.

Heift, T., & Chapelle, C. A. (2012). Language learning through technology. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 555–569). Routledge.

Håkansson Ramberg, M., & Bøhn, H. (2023). Oral skills and learner agency in the foreign language classroom: Reflections on self-regulated learning and feedback from a Nordic perspective. *Nordic Journal of Language Teaching and Learning 11*(1), Tal! Special Issue. https://doi.org/10.46364/njltl.v11i1.1027

ImmerseMe. (n.d.). *VR-based language learning*. https://immerseme.co

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, *23*(4), 1515–1529. https://doi.org/10.1007/s10639-017-9676-0

Johnson, W. L. (2019). Data-driven development and evaluation of Enskill English. *International Journal of Artificial Intelligence in Education, 29*, 425–457. http://doi.org/10.1007/s40593-019- 00182-2

Lantolf, J., Thorne, S. L., & Poehner, M. E. (2014). Sociocultural theory and second language development. In B. VanPatten & J. Williams (Eds.), *Theories in Second Language Acquisition: An Introduction* (pp. 207–226). Taylor and Francis.

Lege, R., & Bonner, E. (2020). Virtual reality in education: The promise, progress, and challenge. *JALT CALL Journal*, *16*(3), 167–180. https://doi.org/10.29140/jaltcall.v16n3.388

Levy, M. (2015). The role of qualitative approaches to research in CALL contexts: Closing in on the learner's experience. *CALICO Journal*, *32*(3), 554–568. Equinox Publishing Ltd.

Li, L. (2017). *New Technologies and Language Learning*. Palgrave.

Lin, T. J., & Lan, Y. J. (2015). Language Learning in Virtual Reality Environments: Past, Present, and Future. *Educational Technology & Society, 18* (4), 486–497. https://www.jstor.org/stable/jeductechsoci.18.4.486

Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching, 51*(3), 285–329. http://dx.doi.org/10.1017/S0261444818000125

Makransky, G., Terkildsen, T., & Mayer, R. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *60*, 225–236.

Mariani, L. (1997). Teacher support and teacher challenge in promoting learner autonomy. *Perspectives: A Journal of TESOL* Italy *23*(2), 5–19.

Meri-Yilan, S. (2019). A constructivist desktop virtual reality-based approach to learning in a higher education institution. In K. Becnel (Ed.), *Emerging Technologies in Virtual Learning Environments*, (pp. 258–283). IGI Global.

Mills, N. (2021). Engagement and immersion in virtual reality narratives. In P. Hiver, A. H. Al-Hoorie & S. Mercer (Eds.). *Student engagement in the language classroom*. Multilingual matters.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16*(1–13). https://doi.org/10.1177/1609406917733847

Papin, K. (2018). Can 360 virtual reality tasks impact L2 willingness to communicate? In P. Taalas, J. Jalkanen, L. Bradley, & S. Thouësny (Eds.) *Future-proof CALL: language learning as exploration and encounters – short papers from EUROCALL 2018* (pp. 243–248). Research-publishing. https://doi.org/10.14705/rpnet.2018.26.844

Parmaxi, A. (2023). Virtual reality in language learning: a systematic review and implications for research and practice. *Interactive Learning Environments*, *31*(1), 172–184. https://doi.org/10.1080/10494820.2020.1765392

Prensky, M. (2016). *Education to better their world: Unleashing the power of 21st-century kids*. Teachers College Press.

Sang, Y., & Hiver, P. (2021). Engagement and companion constructs in language learning: conceptualising learners' involvement in the L2 classroom. In P. Hiver, A. H. Al-Hoorie & S. Mercer (Eds.), *Student engagement in the language classroo*m (pp. 17–37). Multilingual Matters.

Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology*, *109*(3), 431-433. https://doi.org/10.1111/bjop.12305

Soto, J. H. B., Ocampo, D. C. T., del Carmen Beltrán Colón, L., & Oropesa, A. V. (2020). Perceptions of Immerseme virtual reality platform to improve English communicative skills in higher education. *International Journal of Interactive Mobile Technologies*, *14*(7), 4–19. https://doi.org/10.3991/IJIM.V14I07.12181

Sun, Y. H. (2017). Design for CALL – possible synergies between CALL and design for learning. *Computer Assisted Language Learning 30*(6), 575–599. https://doi.org/10.1080/09588221.2017.1329216

Sydorenko, T., Daurio, P., & Thorne, S. L. (2018). Refining pragmatically-appropriate oral communication via computer-simulated conversations. *Computer Assisted Language Learning*, *31*(1–2), 157–180. https://doi.org/10.1080/09588221.2017.1394326

TAL. (n.d.). TAL project. *Learning, Teaching and Assessment of second foreign languages – an alignment study on oral language proficiency in the Swedish school Context.* Lund University. https://tal.blogg.lu.se/in-english/

Thiriau, C. (2017). Teaching speaking in ELT: Global teaching speaking questionnaire.

Cambridge University Press. https://www.cambridge.org/elt/blog/wp-content/uploads/2017/11/Cambridge-globalteaching-speaking-questionnaire-2017.pdf

Thomas, M., Reinders, H., & Warschauer, M. (2012). Contemporary computer-assisted language learning: The role of digital media and incremental change. In M. Thomas, H. Reinders, & M. Warschauer (Eds.), *Contemporary computer-assisted language learning* (pp. 1–12). Bloomsbury Publishing.

Thrasher, T. (2022). The impact of virtual reality on L2 French learners' language anxiety and oral comprehensibility: An exploratory study. *CALICO Journal*, *39*(2). https://doi.org/10.1558/cj.42198

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, Y. F., Petrina, S., & Feng, F. (2017). VILLAGE–Virtual immersive language learning and gaming environment: Immersion and presence. *British Journal of Educational Technology, 48*(2), 431–450. https://doi.org/10.1111/bjet.12388

Wong, G. K., & Notari M. (2018). Exploring immersive language learning using virtual reality. In: M. Spector, B. Lockee & M Childress. (Eds.), *Learning, Design, and Technology* (pp 1–21). Springer. https://doi.org/10.1007/978-3-319-17727-4_144-1

Wood, D. J., Bruner, J. S., & Ross, G. (1976). The Role of Tutoring in Problem Solving. *Journal of Child Psychiatry and Psychology, 17*, 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Yang, F.-C. O., Lo, F.-Y. R., Hsieh, J. C., & Wu, W.-C. V. (2020). Facilitating Communicative Ability of EFL Learners via High-Immersion Virtual Reality. *Journal of Educational Technology & Society*, *23*(1), 30–49. https://www.jstor.org/stable/26915405