



Virtual Reality in Teaching German as a Foreign Language to Learners with Autism: Instructional Proposals

Anna Karagiannidou

Hellenic Open University, Greece

Abstract

Inclusive foreign language education requires differentiated approaches that address learners' individual needs and communication profiles. This paper explores the potential of virtual reality and avatar-based activities to support social communication and vocabulary development in learners with autism spectrum disorder (ASD) studying German as a foreign language. Drawing on relevant literature and the profile of a fifteen-year-old learner with ASD, it presents two instructional proposals. The first draws on the three-dimensional social understanding system (3D-SU) [1] to simulate a visit to a medical practice, using guided interaction and multiple-choice tasks to explore social situations and contextually appropriate responses. The second proposes adapting LIFEisGAME activities [2] to connect facial expression recognition with German vocabulary related to emotions. Tasks include matching expressions to words, modifying avatars, reproducing expressions, and interpreting emotions within short German-language narratives. Both proposals emphasise visual support, structured practice, individualised task design, and teacher guidance. Suggested assessment procedures address learners' responses, vocabulary selection, and representations of facial expressions. The contribution is pedagogical and does not report empirical evaluation of the proposed activities. It outlines possibilities for integrating social and linguistic learning while identifying appropriate interaction modes and the extent of teacher support as priorities for further investigation.

Keywords: virtual reality, German as a foreign language, autism spectrum disorder, inclusive education.

1. Introduction

Inclusive foreign language education requires teachers to align curricular aims with learners' communication profiles, interests and support needs. For learners with autism spectrum disorder (ASD), participation in German language lessons may be influenced by how instructions are presented and how classroom interaction and vocabulary practice are organised. Differentiation therefore involves both access to meaningful communication and the adaptation of instructional content and procedures. Papazoglou discusses the challenges of including learners with autism in the foreign language classroom [3]. In this context, virtual environments and animated characters offer potential resources for designing structured, visually supported learning activities.

This paper presents two instructional proposals informed by the profile of a fifteen-year-old learner with ASD studying German as a foreign language. The first draws on the three-dimensional social understanding system (3D-SU) [1] to propose a simulated visit to a medical practice, combining guided interaction with multiple-choice tasks to explore social situations and contextually appropriate communicative responses. The second draws on LIFEisGAME [2] to propose avatar-based activities that connect facial expression recognition with German vocabulary related to emotions. Both proposals emphasise visual support, structured practice, individualised task design and teacher guidance. Their pedagogical contribution lies in connecting social communication and vocabulary learning within a differentiated approach to German language instruction.

The paper adopts a pedagogical design approach and does not report on the implementation or empirical evaluation of the proposed activities. The learner profile provides a basis for identifying instructional priorities and developing tasks that address specific communication and language learning needs. The German-language examples, activity sequences and assessment procedures are proposed for classroom adaptation. The proposed language content and task sequences should be distinguished from the documented features of the original software. The discussion considers the instructional



rationale, conditions for implementation and limitations of the proposals, with particular attention to appropriate modes of interaction and the extent of teacher support.

2. Pedagogical Rationale

2.1 Virtual Environments and Avatar-Based Learning

Virtual reality and virtual learning environments can provide interactive representations of situations for guided exploration in educational settings. Volioti et al. review the educational uses of virtual environments for children with autism and discuss design considerations, including the control of sensory stimuli and opportunities for individualisation [4]. In the present proposals, an interactive three-dimensional environment and an animated avatar's face serve distinct instructional purposes. The former situates communication within a recognisable social context, whereas the latter allows learners to examine selected visual cues repeatedly. Avatar-based activities presented on a conventional screen should therefore be distinguished from immersive virtual reality experiences.

Cheng, Huang and Yang investigate a three-dimensional immersive virtual environment system designed to support social understanding and social skills in children with ASD [1]. Their study informs the design of the simulated social situation in the first proposal but does not establish the effectiveness of the proposed German-language adaptation. Fernandes et al. describe LIFEisGAME as a facial character animation system designed to support the recognition and expression of emotions [2]. Their work provides the technological and pedagogical basis for the second proposal.

The two proposals share a pedagogical rationale despite their different technological formats. Both are designed to organise practice into manageable steps, provide visual support and incorporate teacher guidance. The first focuses on selecting and using contextually appropriate communicative responses, whereas the second focuses on connecting German emotion words with facial expressions and interpreting these words within short narratives. The potential educational value of both proposals depends on the accessibility of the tasks, their alignment with learning objectives and the opportunities they provide for meaningful language use.

2.2 Individualisation and Communication

Educational accounts of autism describe variability in communication, social interaction and learning profiles [5]. These descriptions should inform the observation of individual learners rather than be treated as a uniform checklist of abilities or difficulties. A learner may recognise written German words while finding an open-ended oral exchange demanding. Similarly, an interest in digital activities does not necessarily indicate a preference for immersive equipment or comfort with all visual and auditory effects. Decisions about task format should therefore be informed by the learner's responses and preferences. In the present proposals, predictable task sequences, concise instructions and a limited number of competing stimuli are central design considerations. The teacher would introduce each activity, demonstrate how to complete the task and adjust the level of support to the learner's needs. Responses produced after modelling would be distinguished from those produced independently. The learner would also have opportunities to request help, take a break or choose an alternative mode of interaction. These options would support differentiated participation while enabling the learner to express needs and make choices within the language lesson.

3. Learner Profile and Scope of the Proposals

The instructional proposals are informed by the profile of a fifteen-year-old learner with ASD attending the second year of lower secondary education at a Greek special vocational secondary school. Her profile indicates an interest in computers, satisfactory reading and writing skills, familiarity with some basic grammatical structures in German and variable engagement in vocabulary activities. It also indicates difficulties in interpreting other people's emotions and limited participation in interactions with classmates.

These characteristics suggest two interrelated instructional priorities: making communicative situations more predictable and connecting familiar written language with supported opportunities for interpretation and response. The learner's interest in computers provides a possible starting point for designing activities, while her existing literacy skills allow the use of short written instructions and sentence frames.



The proposed assessment procedures would distinguish between engagement, linguistic performance and the level of support required for participation.

Both proposals follow a common instructional sequence: defining a specific learning objective, introducing the relevant language, modelling a response, providing guided practice and revisiting the task with reduced support where appropriate. Before implementing the activities, the teacher would assess the learner's understanding of the target vocabulary and ability to use the interface. The examples presented below illustrate how tasks could be designed and adapted to the learner's needs rather than prescribe a fixed instructional programme.

4. First Instructional Proposal: A Visit to a Medical Practice

4.1 Aims and Preparation

The first proposal draws on the rationale of 3D-SU [1] to design a simulated visit to a medical practice. The social communication objectives are to support the learner in recognising the sequence of the visit, interpreting simple situational cues and selecting contextually appropriate responses. The language objectives are to develop her understanding of short questions and her ability to use expressions for greeting, waiting, asking for information and requesting assistance. Before introducing the activity, the teacher would establish whether the medical setting is relevant to the learner's needs and whether she feels comfortable exploring it.

An introductory vocabulary set could include Arzt and Ärztin (doctor), Wartezimmer (waiting room), warten (to wait) and Hilfe (help). Useful expressions include "Guten Tag" (Good day), "Wann bin ich dran?" (When is it my turn?) and "Ich brauche Hilfe" (I need help). Before introducing the simulation, the teacher would present these items through written prompts, images and brief demonstrations of their use. A visual sequence could illustrate the stages of the visit: arrival, greeting, waiting, consultation and departure. The initial presentation would be limited to the language required for the immediate task.

Implementation would require access to a suitable virtual environment and an interaction method accessible to the learner. The simulated setting and German-language prompts constitute proposed instructional adaptations; they are not presented as existing features of the published 3D-SU system. Access to the software, options for language customisation and equipment requirements would need to be verified before classroom implementation. If an immersive format were inaccessible or uncomfortable for the learner, a screen-based alternative could preserve the instructional sequence.

4.2 Guided Interaction and Feedback

The teacher would first demonstrate how to navigate the virtual environment and select a response, without assessing the learner's performance during this introductory stage. At the first decision point, the learner would see people waiting and read or hear the question "Was machst du im Wartezimmer?" (What do you do in the waiting room?). Two response options could be "Ich warte, bis ich dran bin" (I wait until it is my turn) and "Ich gehe sofort ins Sprechzimmer" (I go straight into the consulting room). The teacher would relate the selected response to the situation depicted and explain why waiting is appropriate in this context.

A second decision point would provide an opportunity to request assistance. For example, the learner could select "Entschuldigung, ich brauche Hilfe" (Excuse me, I need help). This task would support both the recognition of a familiar routine and the expression of an individual need. The prompts would avoid suggesting that waiting silently is the only acceptable response in every situation. If the circumstances change, asking a question or seeking help may be appropriate. The teacher would therefore explain how communicative choices depend on the context.

Feedback would be immediate, concise and adapted to the learner's needs. Following a contextually appropriate response, the screen could display "Richtig. Du wartest, bis du dran bist" (Correct. You wait until it is your turn). Following an inappropriate response, it could display "Schau noch einmal. Die anderen warten auch" (Look again. The others are waiting too). The learner would be offered another attempt and, where necessary, a model response. Sound effects would be optional, and the number of attempts would be adjusted according to the learner's responses and comfort.

Following the guided response-selection activities, the teacher could invite the learner to use the same expressions in a brief classroom role-play. Written prompts or a visual sequence would remain available as needed. Observing the learner's responses in both the virtual task and the classroom exchange would provide opportunities to examine her understanding and use of the target language across the



two formats. The immediate objective would be participation in a manageable German-language exchange, with the type and level of assistance recorded to inform subsequent teaching.

5. Second Instructional Proposal Emotion Vocabulary with Avatars

5.1 Aims and Task Progression

The second proposal draws on the facial animation approach of LIFEisGAME [2] and organises vocabulary practice into four stages. It focuses on a small set of German adjectives describing emotions, initially *fröhlich* (happy or cheerful), *ärgerlich* (angry or annoyed) and *traurig* (sad). The sequence progresses from supported recognition to the construction of facial expressions, optional reproduction of expressions and contextual interpretation within a short narrative. The German-language prompts and narrative are proposed instructional adaptations; their availability within the original software is not assumed.

In the first stage, the learner would view an avatar displaying a clearly distinguishable facial expression and choose between two written words, such as *fröhlich* and *ärgerlich*. Greek translations could initially support comprehension. The teacher would pronounce the selected word and draw attention to relevant visual features, while explaining that facial expressions do not always provide a reliable indication of a person's feelings. Different avatar faces and varied positions of the response options would subsequently be introduced to reduce reliance on a single image or the location of the correct answer.

In the second stage, the learner would select an avatar and modify its facial expression in response to a written German adjective. For example, *traurig* would appear without a Greek translation, and the learner would use the available controls to construct an expression associated with sadness. The teacher could ask "Wie fühlt sich die Person?" (How does the person feel?) and provide the sentence frame "Die Person ist ..." (The person is ...), which the learner could complete with *traurig*. Word recognition, sentence completion and use of the interface would be assessed separately, so that difficulty operating a control would not automatically be interpreted as a lack of understanding of the adjective.

In the third stage, the learner could compare the avatar's expression with a model provided by the teacher or, if she felt comfortable, reproduce an expression herself. Taking a photograph of her own expression would be optional; teacher-provided images or illustrations could serve as alternatives. Any use of identifiable photographs would require appropriate consent and arrangements for their storage and use. The language objective would be to connect emotion words with visual representations and discuss these associations. Reproducing a particular facial expression would not be a prerequisite for successful participation.

5.2 Emotion Words in a Short German Narrative

The fourth stage would place the target vocabulary within a short narrative set in a school. A possible text would be: "Irene ist in der Schule. Ihre Freundin kommt. Irene ist *fröhlich*. Später findet sie ihr Heft nicht. Sie ist *traurig*. Ein Mitschüler nimmt ihren Stift, ohne zu fragen. Irene ist *ärgerlich*. Sie sagt: 'Bitte frag mich zuerst.'" The narrative links each emotion adjective to a specific event, while the final utterance introduces a simple expression for communicating a personal boundary.

After each event, the teacher would ask "Wie fühlt sich Irene?" (How does Irene feel?). The learner could select an adjective, complete the sentence frame "Irene ist ..." or create a corresponding facial expression on an avatar. A follow-up task could invite her to identify the event associated with Irene's stated emotion by choosing between two brief descriptions or images. Since the narrative explicitly names the emotions, correct responses at this stage would indicate comprehension supported by explicit textual cues, rather than independent inference. A subsequent version could omit an emotion adjective and invite the learner to select a plausible word on the basis of the narrative context.

Further practice could vary the characters, events or order of the response options while retaining the target vocabulary. The teacher would explain that people may experience different emotions in similar situations. Facial representations would serve as instructional cues within the narrative rather than as definitive indicators of a person's emotional state. In tasks where the emotion is not explicitly stated, the teacher would discuss alternative interpretations that the learner could support with reference to the



context, rather than automatically treating responses that differ from a predetermined answer as incorrect.

6. Assessment and Conditions for Implementation

Assessment would be formative and aligned with the stated learning objectives. Before introducing the activities, the teacher would establish which target words and expressions the learner already understands. During the simulated visit to a medical practice, observations would focus on the learner's understanding of the situation, selection of a contextually appropriate response and use of the target German expressions. During the avatar-based activities, assessment would distinguish between word–image matching, word recognition without translation, supported sentence production and interpretation of emotion vocabulary in context.

For each task, a brief observation record could document the item presented, the learner's initial and subsequent responses, and the type and level of assistance provided. Support could include repetition of the instruction, a visual cue, a reduced set of response options or a model response. Initial responses would be distinguished from those produced after feedback, since successful task completion does not necessarily indicate independent understanding. Revisiting the target language after an interval and using varied examples could help assess retention and the ability to apply knowledge across tasks. Classroom role-play or discussion could provide additional opportunities to observe language use beyond the original digital activity.

The learner's representations of facial expressions would be considered alongside her vocabulary choices and explanations of the narrative context. Visual similarity to a model would not be treated as a general measure of emotional understanding. Likewise, attention to the screen and enjoyment of the activity would be recorded separately from language performance. The learner's comments, preferred response mode and requests for breaks would help the teacher determine whether the activity format remains accessible and appropriate. Assessment would therefore guide subsequent instructional decisions rather than reduce different aspects of performance to a single score.

Practical implementation would require suitable equipment, accessible controls and language, and adequate teacher preparation. The teacher would review all content in advance, introduce the interface gradually and adjust visual complexity and sound to the learner's needs. Screen-based alternatives would be available where appropriate. A future empirical evaluation should specify the duration of instruction, the learner's initial knowledge of the target language, the task conditions and the observation procedures. It should also examine whether accurate responses are maintained as support is reduced and distinguish gains in language knowledge and use from improvements attributable primarily to familiarity with the interface or repeated exposure to the same items.

7. Discussion and Conclusions

The two instructional proposals connect social communication and language learning objectives through different forms of supported practice. The simulated visit to a medical practice links situational understanding with expressions for participation, requesting information and seeking assistance. The avatar-based sequence connects German emotion adjectives with facial representations and short narratives. Together, the proposals illustrate possible ways of incorporating virtual environments and avatar-based activities into differentiated foreign language teaching. Both emphasise clearly defined tasks and responsive teacher guidance, with the choice of technology informed by the learner's needs and the intended learning objectives.

Several limitations define the scope of this contribution. The proposals have not been empirically evaluated, and the individual learner profile does not provide a basis for generalising their suitability to other learners with ASD. The selected literature offers a rationale for exploring virtual environments and facial animation but does not establish the effectiveness of the specific German-language activities proposed here. Access to suitable software and the feasibility of customisation would need to be established before implementation. Furthermore, selecting an appropriate response in a virtual task or identifying a stylised facial expression would not, in itself, demonstrate independent communication or an understanding of emotions in everyday situations.

Further research should examine the accessibility of different interaction modes, the type and level of teacher support required, and the extent to which learners can retain and use the target language beyond the practised tasks. The principal pedagogical contribution of this paper is an explicit connection between an individual learner profile, specific learning objectives, supported activities and proposed



assessment procedures. Virtual reality and avatar-based resources may offer useful opportunities for integrating social communication and vocabulary practice when their use is guided by individual needs and instructional aims. Establishing their educational value would require systematic evaluation of learning outcomes, including retention and transfer, alongside consideration of the learner's experience of participation.

REFERENCES

- [1] Cheng Y., Huang C.-L., Yang C.-S., "Using a 3D Immersive Virtual Environment System to Enhance Social Understanding and Social Skills for Children With Autism Spectrum Disorders", *Focus on Autism and Other Developmental Disabilities*, 30(4), 2015, pp. 222–236. <https://doi.org/10.1177/1088357615583473>
- [2] Fernandes T., Alves S., Miranda J., Queirós C., Orvalho V., "LIFEisGAME: A Facial Character Animation System to Help Recognize Facial Expressions", in Cruz-Cunha M. M., Varajão J., Powell P., Martinho R. (eds.), *ENTERprise Information Systems, Communications in Computer and Information Science*, Vol. 221, Berlin and Heidelberg, Springer, 2011, pp. 423–432. https://doi.org/10.1007/978-3-642-24352-3_44
- [3] Παπάζογλου Δ., «Ο Μαθητής με Αυτισμό στην Ξενόγλωσση Τάξη» [The Learner with Autism in the Foreign Language Classroom], in Γεωργογιάννης Π. (ed.), *Πρακτικά 19ου Διεθνούς Συνεδρίου «Εκπαίδευση και Ετερότητα» [Proceedings of the 19th International Conference Education and Diversity]*, Vol. 1, Πάτρα, Ελεύθερο Πανεπιστήμιο Πολιτών, 2015, pp. 424–436. [In Greek].
- [4] Βολιώτη Χ., Τσιάτσος Θ., Μαυροπούλου Σ., Καραγιαννίδης Χ., «Εικονικά Περιβάλλοντα Μάθησης για Παιδιά με Αυτισμό: Επισκόπηση Πεδίου και Προτάσεις Σχεδιασμού» [Virtual Learning Environments for Children with Autism: A Review of the Field and Design Proposals], in Λαδιάς Α., Μικρόπουλος Α., Παναγιωτακόπουλος Χ., Παρασκευά Φ., Πιντέλας Π., Πολίτης Π., Ρετάλης Σ., Σάμψων Δ., Φαχαντίδης Ν., Χαλκίδης Α. (eds.), *Πρακτικά 3ου Πανελληνίου Συνεδρίου «Ένταξη των ΤΠΕ στην Εκπαιδευτική Διαδικασία» [Proceedings of the 3rd Panhellenic Conference Integration of ICT in the Educational Process]*, Πειραιάς, Ελληνική Επιστημονική Ένωση Τεχνολογιών Πληροφορίας και Επικοινωνιών στην Εκπαίδευση, 2013, pp. 169–176. <https://www.etpe.gr/wp-content/uploads/pdfs/etpe2064.pdf> [In Greek].
- [5] Μαυροπούλου Σ., «Αποτελεσματικές εκπαιδευτικές προσεγγίσεις και διδακτικές στρατηγικές για τα παιδιά στο φάσμα του αυτισμού» [Effective Educational Approaches and Teaching Strategies for Children on the Autism Spectrum], in Παντελιάδου Σ., Αργυρόπουλος Β. (eds.), *Ειδική αγωγή: Από την έρευνα στη διδακτική πράξη [Special Education: From Research to Teaching Practice]*, Αθήνα, Πεδίο, 2011, pp. 83–134. [In Greek].