Reframing Speech Therapy Through Design: Gamification, Emotion, and Embodied Interaction

Patrizia Marti
Università degli Studi di Siena
patrizia.marti@unisi.it
ORCID 0000-0002-2448-8747

Andrea Brogi
Logopedista e Psicologo
brogilogopsicologo@gmail.com
ORCID 0009-0001-0449-9924

Alessia Nicoletta Marino
Università degli Studi di Siena / Università degli Studi della Campania Luigi Vanvitelli
a.marino13@student.unisi.it
ORCID 0009-0004-7105-5305

Abstract
Gamification and conversational technologies are increasingly explored in paediatric rehabilitation to support children's engagement with therapeutic activities. Within speech therapy, however, the integration of *Embodied Conversational Agents* (ECAs) remains underexplored from a design perspective, particularly with regard to children's emotional experience, motivation and interaction dynamics. This paper investigates how observations of children's responses to errors and corrective feedback can inform the design of ECAs for paediatric speech therapy. Adopting a human-centred and design-driven approach, the research combines a review of existing gamified therapeutic systems, non-participant observation of clinical speech therapy sessions, and interviews with therapists. The findings highlight recurrent differences in children's responses to errors, the importance of relational sensitivity in therapeutic interactions, and challenges in extending therapeutic support to home-based practice. These insights were translated into design principles concerning non-judgmental error mediation, adaptive scaffolding, narrative reframing, and calibrated relational embodiment. They informed the development of *Ziza's Words*, a narrative-based game featuring an embodied agent designed as a playful companion rather than an authoritative evaluator. A preliminary formative pilot with three children provided initial insights into engagement, emotional responses, and interaction affordances. The contribution lies in articulating preliminary design propositions for emotionally sensitive ECAs in paediatric rehabilitation.

Keywords
Embodied conversational Agent
Speech therapy
Paediatric rehabilitation
Gamification
Emotional interaction

Copyright © 2026
diid disegno industriale industrial design
CC BY-NC-SA

Embodied Conversational Agents in paediatric speech therapy

Primary language disorders are among the most common difficulties in childhood, mainly affecting sound production; in some cases, they also impair comprehension. They cannot be classified as a single diagnostic category, but rather as a heterogeneous set of conditions that can vary in nature and severity. For these reasons, it is essential to identify language disorders early, between the ages of 3 and 4 (Rinaldi et al., 2021).

Gamification techniques are increasingly used in paediatric speech therapy to enhance patients' motivation and engagement throughout the therapeutic process. Teverini and colleagues (2025) describe it as a design-driven strategy aimed at shaping patient engagement, motivation and sustained participation through interactive systems. Rather than treating gamification solely as a behavioural incentive, it is presented as a tool capable of structuring the therapeutic experience through carefully designed feedback mechanisms, goals, narratives and progression systems. This approach can be employed both during assessment activities and in clinical practice itself. For example, ArticuMotion embeds speech assessment tasks within interactive narratives and allows therapists to customise task parameters (Alsebayel et al., 2024).

Gamified speech therapy systems employ different mechanisms to sustain engagement, including narrative contexts, rewards, character customisation, progression and adaptive difficulty. For example, Apraxia World uses extrinsic rewards and character customisation (Hair et al., 2018), while LogoHra and Flappy Voice allow therapists to customise challenges across different difficulty levels and to monitor patients' activities (Kováč & Ondáš, 2025; Lan et al., 2014). Taken together, these approaches suggest that gamification can transform repetitive therapeutic exercises into more engaging activities; however, they also show that engagement depends on how therapeutic goals, game mechanics, feedback and progression are integrated into the overall experience.

Conversational technologies represent a promising extension of these approaches (Antunes et al., 2023). They can support communication-based activities and potentially extend therapeutic exercises to the home under parental supervision, while reducing therapists' workload (Oliveira et al., 2025; Saeedi et al., 2022).

In designing these technologies, as children's cognitive, physical and socio-emotional capacities are still developing, Child-Computer Interaction highlights the need for technology that adapts to their specific needs (Read & Markopoulos, 2013) as well as to intersectional identities, diverse abilities and linguistic and cultural differences (Garg et al., 2022). The use of non-embodied conversational agents, such as voice assistants for telemedicine systems (Qiu & Abdullah, 2021) and dialog-based conversational agents, has been widely investigated, showing positive impacts in supporting individuals with communication disorders (Khan et al., 2025; Imaezue & Marampelly, 2025).

Embodiment may further influence children's engagement with conversational agents (Spitale et al., 2020), which can take

different forms, including bio-inspired, functional and mechanical embodiments (Cassell, 2000). However, systematic applications of Embodied Conversational Agents in speech therapy remain limited, underscoring the need for further investigation into the role of embodiment in these therapeutic contexts. The adoption of new technologies in rehabilitative contexts should not consider therapeutic goals alone, but also the emotional and relational needs of children that emerge throughout the rehabilitation process, together with the ethical issues concerning personal data protection, emotional attachment and dependency, which may represent a significant concern. Ethical concerns also include the protection of children's personal and health data, as well as transparency and data minimisation.

Technical and interactional difficulties can also elicit frustration in young users (Saeedi et al., 2022). Throughout the rehabilitation process, the child must in fact manage a significant emotional load, related not only to the technological aspects of the interaction, but also to the physiological trial-and-error process that characterises the rehabilitative pathway. Embodied agents designed to display empathy, encouragement, or companionship may foster strong affective bonds that can support engagement but also raise concerns about emotional reliance and reduced human interaction. Designers must therefore carefully calibrate social cues and avoid deceptive anthropomorphism.

Research through Design conceptualises the design practice and the resulting artefacts as means through which researchers can investigate open-ended problems and generate situated knowledge (Zimmerman et al., 2007). Similarly, Dalsgaard (2010) emphasises the role of design experimentation in producing knowledge through the making and use of artefacts, while constructive design research considers designed artefacts as vehicles for exploring, materialising and communicating design knowledge (Koskinen et al., 2011). This perspective is particularly relevant to ECAs in healthcare, where interaction qualities such as embodiment, feedback, emotional expression and relational distance emerge through situated encounters rather than being fully defined in advance. From this perspective, the design of an ECA is not merely a matter of implementing therapeutic functions, but also a means of investigating how aspects of therapeutic practice can be translated into interactional and relational design strategies.

Despite these developments, limited attention has been devoted to how children's emotional and relational responses within actual therapeutic interactions can inform the design of ECAs for paediatric speech therapy. However, less attention has been paid to translating the therapist's situated strategies for managing errors, frustration, encouragement and engagement into explicit design principles for ECAs. This study addresses this gap by connecting observations of clinical speech therapy interactions with a design process aimed at articulating and materialising design principles for emotionally sensitive conversational interactions in paediatric rehabilitation.

Objectives and methodology

Through an empirical investigation, this paper aims to answer the following research question:

RQ: How can observations of children's emotional responses to errors and corrective feedback inform design principles for embodied conversational agents supporting home-based paediatric speech therapy?

The objective is not to assess the clinical effectiveness of an ECA, but to investigate how insights from therapeutic practice can be translated into design strategies for emotionally sensitive conversational interactions. Accordingly, the study distinguishes between four interconnected stages: (1) empirical investigation of therapist–child interactions; (2) qualitative analysis of emotional and relational patterns; (3) translation of these findings into design principles and an ECA prototype; and (4) a preliminary formative pilot exploring the usability and experiential qualities of the resulting design.

The study therefore adopts a qualitative, design-driven research approach, in which the prototype functions as an artefact through which research insights are materialised and explored.

The research adopts a human-centred and research-through-design orientation, using an iterative design process to translate empirical findings into design propositions and a prototype.

Data collection and analysis

An empathisation phase was conducted to gain an in-depth understanding of patients' needs, behaviours and attitudes towards speech therapy. This phase began with desk-based research on established therapeutic protocols and a benchmarking analysis of existing digital solutions for speech therapy. These preliminary activities informed an intensive period of non-participant observation of therapist–patient interactions during clinical sessions. The observations aimed to capture not only the structure of therapeutic activities and clinical strategies, but also the emotional processes emerging during interaction.

Tab. I
Observation grid: patient information, assessments of morphosyntactic competence and language comprehension and therapy activity.

Patient code:		**Language Expression**						
Gender		*Pre-syntactic phase* (Telegraphic word sequences, lacking free morphemes)	☐					
Age Range		*Early syntactic phase* (Simple, often incomplete utterances; complex sentences appear without sentence connectives and morphemes)	☐					
Type of Exercises		*Nuclear sentence completion phase* (Many nuclear sentences with free morphemes and expanded sentences with core expansion and adverbials; complex sentences become more diversified)	☐					
		Complex combinatorial structures phase (Sentences become grammatically correct)	☐					
Therapy Phases								
Welcome		**Exercise 1**		**Exercise 2**		**Exercise 3**		**Feedback with caregiver**
Patient PrEmo		**Patient PrEmo**		**Patient PrEmo**		**Patient PrEmo**		**Patient PrEmo**
Other emotions:		Other emotions:		Other emotions:		Other emotions:		Other emotions:
Therapist PrEmo		**Therapist PrEmo**		**Therapist PrEmo**		**Therapist PrEmo**		**Therapist PrEmo**
What elicited the emotional state?		What elicited the emotional state?		What elicited the emotional state?		What elicited the emotional state?		What elicited the emotional state?
Notes:		Notes:		Notes:		Notes:		Notes:

diid No. 89 — 2026
Doi: 10.30682/diid8926i

For data collection, a structured observation grid was developed and used in an anonymous manner. The grid included general patient information (sex, age range, date of observation and primary exercise types), assessments of morphosyntactic competence and language comprehension based on the Bortolini scale (Perrotta & Rustici, 2023) and a therapy activity scheme analysed using the PrEmo2 framework (Desmet, 2018) to identify and associate emotional responses with specific therapeutic tasks.

Data analysis

The observational data were analysed qualitatively, focusing on episodes in which children encountered errors, received corrective feedback, and subsequently changed their behaviour or emotional state.

Observation-grid entries and field notes were organised according to the phases of the therapeutic session and examined comparatively across participants. Emotional responses were documented using the PrEmo2 framework, which provided a structured vocabulary for recording observed affective responses.

The analysis followed an inductive clustering process in which recurring patterns of response to errors and corrective feedback were compared across sessions. This process led to the identification of three recurrent behavioural-emotional profiles, which emerged as a preliminary clustering from the synthesis of empathy-driven methods:

- children who view errors as learning-oriented response
- children who experience errors as failure-sensitive response
- children who exhibit oppositional/frustration response to errors

These profiles are observed interactional profiles rather than clinical categories or psychological typologies.

In total, twenty-three therapeutic sessions were observed, involving fifteen paediatric patients (eleven male and four female): one aged 4–5 years, twelve aged 7–8 years and two aged 12–13 years. Participants were selected by the speech therapist, who invited one of the co-authors of this study to act as a non-participant observer during clinical sessions.

Prior to the observations, informed consent was obtained from the parents, who were informed that the presence of the researcher would not interfere with or alter the therapeutic pathway.

Systematic observation revealed a recurring structure within the 45-minute sessions, consisting of:

- an initial welcome phase;
- a first exercise;
- a phase of structured play;
- a correction phase or second exercise and
- a final feedback moment involving the parent, often accompanied by the assignment of home-based exercises.

Beyond the formal organisation of therapy, observations focused on therapist–child interaction. Children's responses to errors and corrective feedback varied and were closely intertwined with engagement. The therapist continuously adapted exercises, feedback and interaction strategies to these responses, revealing tacit relational strategies relevant to the design of emotionally sensitive technologies.

Semi-structured interviews

The observational phase was complemented by 10 semi-structured interviews with 5 speech and language therapists, conducted in a first phase between March and July 2025, during which detailed notes were taken, and in a second phase between May 2026 and June 2026, conducted online via Teams. With participants' consent, interviews conducted in the second phase were audio and video recorded and automatically transcribed with an AI tool, then manually checked and corrected, while interviews conducted in the first phase were documented through detailed notes. Interviews lasted approximately 30 minutes.

The interview protocol addressed three main topics: clinical practice (professional background, commonly treated disorders, therapeutic materials and tools, and home-practice assignments); incentives and error management (motivation strategies, reward systems, and children's emotional responses to errors); and the relationship with technology (current use of digital tools, technological support, and adaptation of therapeutic activities).

Interview material was analysed qualitatively through thematic analysis, focusing on recurring challenges and strategies relevant to the design of home-based therapeutic interactions.

The thematic analysis revealed key challenges: low compliance with home-based exercises, difficulties in monitoring activities carried out between sessions and the lack of technological tools capable of adapting therapeutic content to the emotional states and personality traits of individual patients.

The interviews further highlighted therapist–patient interaction as central to care, particularly the role of attention, trust and sensitive verbal and non-verbal communication (Busch et al., 2019). Therapists described errors as opportunities for learning rather than failures, using encouragement, guided reflection and trust-building to support children's engagement.

Co-design and translation into design decisions

The empirical findings were subsequently discussed and translated into design decisions through two co-design sessions, lasting 30 minutes and involving the design researchers and a speech therapist. The sessions focused on the construction of two scenarios: the first was a general overview aimed at illustrating the technology and envisioning its possible applications, while the second was more specific, involving a trial of a conversational agent.

The co-design process was used to translate clinical knowledge into interaction requirements and design propositions, with particular attention to error responses, encouragement, and the visual and vocal embodiment of the ECA.

Design decisions were iteratively reviewed against the observational findings and clinical requirements.

The empirical findings were translated into a set of design principles that guided the development of the prototype. Rather than treating the observed emotional responses as individual character-

diid No. 89 — 2026
Doi: 10.30682/diid8926i

istics of the children, the analysis interpreted them as interactional conditions that may require different forms of support from the technological system. This led to a design approach in which emotional sensitivity became a core design variable alongside therapeutic functionality and engagement.

Six main design directions emerged from the empirical investigation. First, because errors elicited different emotional responses across children, the system should mediate errors through supportive and non-judgmental feedback rather than explicit evaluation. Second, because the therapist continuously adjusted the difficulty and interaction strategy according to each child's response, the game should incorporate adaptive scaffolding capable of modulating task difficulty and feedback. Third, because narrative and play helped transform repetitive exercises into meaningful activities, therapeutic tasks should be embedded within a coherent narrative context rather than presented as isolated exercises. Fourth, because the therapist's relational role was central to sustaining trust and engagement, the embodied agent should be designed as a supportive companion rather than an authoritative evaluator, with its visual, vocal, and behavioural characteristics carefully calibrated to establish an appropriate relational distance. Fifth, because children differed in attention and engagement, the interaction should accommodate different modes and levels of engagement. Sixth, because therapists need to assign and monitor home exercises, the system should support autonomous home practice and provide visibility into activities performed outside clinical sessions.

Empirical finding	Design implication	Design principle	Prototype implementation
Children respond differently to errors and corrective feedback	Feedback should be sensitive to emotional responses and avoid increasing performance pressure	**Non-judgmental error mediation**	Encouraging prompts, actionable suggestions, and avoidance of explicit negative evaluation
Therapists continuously modulate exercises and feedback according to children's responses	The system should be able to adjust the level of challenge	**Adaptive scaffolding**	Reduction/increase of task difficulty according to performance
Narrative and play can transform repetitive exercises into meaningful activities	Therapeutic tasks should be embedded in a meaningful context	**Narrative reframing**	Ziza's story, missions, magical honey, and game progression
The therapist's relational sensitivity supports trust and engagement	The agent should complement rather than reproduce an authoritative therapeutic role	**Supportive relational role**	Pierina designed as a playful companion rather than evaluator
Embodied agents can evoke social expectations and emotional responses	Embodiment should be calibrated according to the intended relationship	**Calibrated relational embodiment**	Toy-like appearance, visible mechanical components, synthetic child-like voice, avoidance of excessive realism
Children differ in attention and engagement	The interaction should accommodate different modes and levels of engagement	**Flexible engagement**	Choice of story access and different game activities
Therapists need to assign exercises at home between sessions and to monitor the outcomes achieved	The system should support home practice and provide therapists with visibility into children's performance outside sessions	**Autonomous home practice**	Tracking of the activities at home

Tab. II
Design principles and prototyping hints emerged from the co-design sessions.

These principles provided the design rationale for the subsequent development of Ziza's Words and Pierina. The relationship between empirical findings, design principles, and their implementation in the prototype is summarised in Tab. II.

Ziza's Words: a gamified therapeutic approach to speech therapy

A video game was developed integrating a conversational robotic agent, named Pierina, designed to support phoneme perception and production through playful and narrative-driven interactions. Rather than presenting speech exercises as isolated or repetitive tasks, the game embeds therapeutic activities within an adventure-based structure grounded in storytelling. Ziza's Words embeds phoneme perception and production exercises within an adventure-based narrative centred on Ziza, a small bee facing a series of challenges. Children help Ziza progress through the story by completing therapeutic activities, transforming repetitive exercises into meaningful actions within the game world. This narrative reframing provides coherence and purpose while reducing the emphasis on correction and performance. Game dynamics such as progression, challenge and achievement are carefully calibrated to sustain motivation while avoiding performance pressure. Within this framework, Pierina guides the child through the activities and provides encouragement without explicitly evaluating performance. The principle of calibrated relational embodiment informed the visual and behavioural design of Pierina, a toy robot that guides the child through the activities, offers encouragement and defers judgement during task execution, thereby reducing the anxiety associated with errors and fostering a safe learning environment.

The conversational agent was developed using Algho, a no-code/low-code AI platform developed by QuestIT s.r.l., enabling speech-to-text and text-to-speech interaction. The design focused on visual appearance, expressive behaviour and interaction style, selectively incorporating human-like traits while avoiding excessive realism (Carolus et al., 2022; Forlizzi et al., 2007).

Overall, the integration of narrative-driven gameplay, expressive embodiment and non-judgmental coaching positions the embodied conversational agent as a mediating design artefact that supports therapeutic engagement, emotional regulation and sustained participation, highlighting the potential of design-led approaches in reimagining paediatric speech therapy.

Research on human–agent interaction suggests that users tend to apply social norms to conversational agents (Reeves & Nass, 1996). Accordingly, Pierina was designed as a playful companion rather than an authority figure, favouring support and collaboration over evaluation.

The visual design of the agent deliberately avoids associations with adult professional roles, such as doctors, therapists, or parents, which could implicitly evoke authority and judgement. Instead, Pierina was conceived as a toy-like robot characterised by visible mechanical components and exposed circuits.

This aesthetic strategy seeks to balance approachability and engagement with a clear emphasis on the agent's artificial and non-human nature (Murtarelli et al., 2021). By foregrounding its inanimate qualities, the design mitigates the risk of emotional over-identification or dependency, while still maintaining an appealing and friendly presence. The interaction environment was intentionally kept minimal and visually uncluttered in order to reduce cognitive load and avoid distracting elements that might interfere with task execution.

Beyond visual appearance, careful attention was devoted to the agent's vocal and communicative behaviour. Pierina employs simple, informal language that mirrors children's everyday speech patterns, supporting accessibility and relational proximity.

The vocal design followed the same principles (Poushneh, 2021): a neural synthetic voice was selected to resemble that of a young girl, conveying warmth and familiarity without striving for excessive realism. Together, these design decisions contribute to a coherent interaction style in which embodiment, language and voice collectively support engagement, emotional safety and sustained participation within the gamified therapeutic experience.

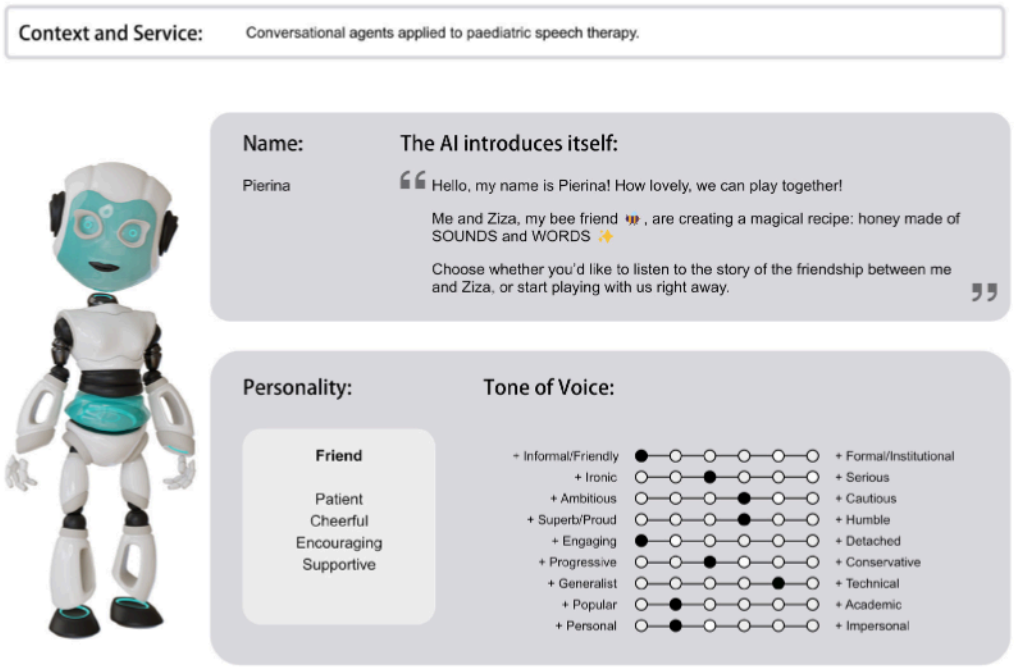


Fig. 1
Overview of Pierina's persona: introductory dialogue, core personality traits and the intended tone of voice, to guide consistent character design.

Game structure

The game experience is embedded within a narrative framework centred on the overarching goal of helping the bee Ziza overcome a series of challenges. At the beginning of the interaction, the embodied conversational agent Pierina introduces herself and presents Ziza, offering the child the option to start playing immediately or to listen to the introductory story, either at the first access or at a later stage. This flexible entry point accommodates different levels of attention and familiarity, allowing children to engage with the experience at their own pace.

The gameplay is organised into four main phases:

- an engagement phase,
- the introduction of the target phoneme through an explanatory video,
- a phoneme production activity,
- a phoneme perception (listening) activity.

The narrative introduction establishes the fictional world and frames therapeutic tasks as missions, providing a meaningful transition into the rehabilitative activity.

Within the story, children are asked to contribute by correctly responding to exercises in order to produce “magical honey” made of sounds and words, which symbolically represents progress within the game world.

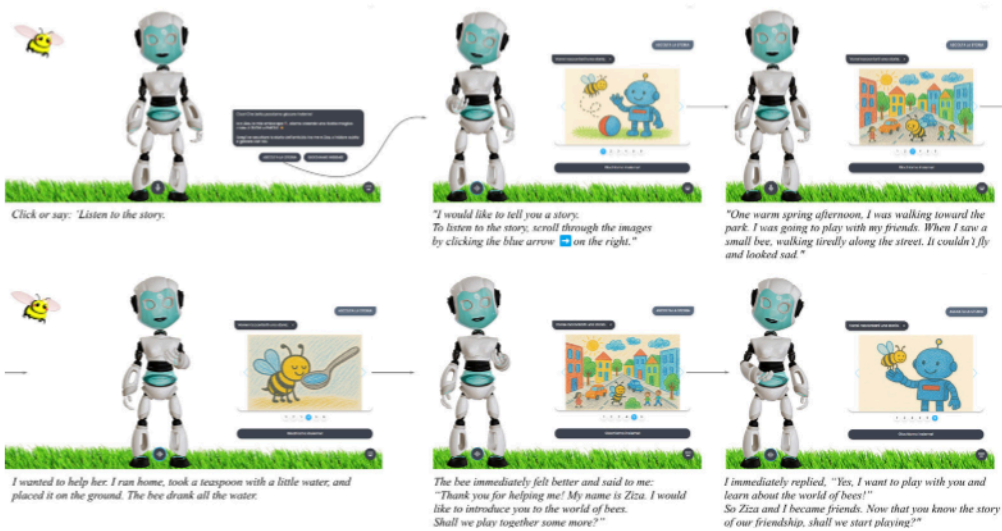


Fig. 2
Game startup screens: intro narrative where Ziza the bee meets Pierina, framing exercises within a meaningful storyline context.

Before accessing the minigames, children are invited to select the phoneme on which they wish to work. The selected phoneme is introduced through a short explanatory video designed to support comprehension and anticipation.

diid No. 89 — 2026
Doi: 10.30682/diid8926i

Children can then choose between two game modes: listening (phoneme perception) or production.

In the listening mode, the target phoneme is introduced through an onomatopoeic and visual association embodied by an animal character—for example, a snake whose sound evokes the phoneme /s/. The activity is structured as a collection game in which the child identifies and selects flowers containing images of words in which the phoneme appears in initial or medial position. Correct responses result in the collection of the flower, while errors trigger encouraging prompts that invite the child to try again without negative judgement.



Fig. 3
Perception-based exercise: identifying the target sound /S/ within a word, associated here with the image of a sun.

In the production mode, the child is asked to recognise an image depicted in a drawing created by the bees and to pronounce the corresponding word aloud. As in the listening activity, the target phoneme may appear in either initial or medial position.

At the end of each activity, the system provides positive reinforcement, modulated according to performance, with the aim of supporting motivation while avoiding performance anxiety.

Explorative formative pilot study

The pilot was conceived as a formative and exploratory evaluation, rather than as a validation study. Its purpose was to explore how children with different previously observed interactional profiles responded to the prototype and to identify design issues requiring further iteration. Given the small sample (N = 3), the observations are not intended to establish therapeutic effectiveness, generalise behavioural patterns, or compare outcomes across participants.

For the preliminary pilot study, the intervention focused on the phoneme /s/.

Three children—two male and one female—were selected by the speech therapist according to predefined inclusion and exclusion criteria that he defined at the beginning of the therapy. Inclusion criteria comprised children aged 5–8 years diagnosed with primary language disorder who had previously worked on the /s/ phoneme during therapy sessions.

Exclusion criteria included children who were unlikely to tolerate interaction with the embodied conversational agent and those who had not yet begun therapy targeting the /s/ phoneme. The three participants were purposively selected to explore the prototype with children representing the three interactional patterns identified during the observational phase and based based on the potential benefit of additional practice on the target phoneme. These profiles were used as design-oriented sampling categories, not as clinical or psychological classifications.

The activity was observed by a co-author of this paper who acted as a non-participant observer.

At the beginning of the study, the therapist introduced the game by narrating the story of the friendship between Ziza and Pierina. The story was subsequently retold by the conversational agent through spoken narration accompanied by scrolling images displayed on a computer screen, supporting comprehension and engagement.

The phoneme perception activity was implemented as a collection game in which participants were required to identify and select flowers containing images of words featuring the /s/ phoneme in initial position.

The therapist evaluated interest (spontaneous attraction towards the activity or its components driven by curiosity and immediate enjoyment), attention during story listening, and the appropriateness of the narrative plot and language; emotional responses were documented using PrEmo2.

The first participant showed immediate interest, sustained attention and joy throughout the narrative. The second initially showed limited interest but increased attention once the story began; observed emotions included hope and diffidence. The third showed high initial interest but progressively disengaged, with joy and boredom predominating. These observations indicate variability in narrative engagement across interactional profiles rather than a uniform motivational effect.

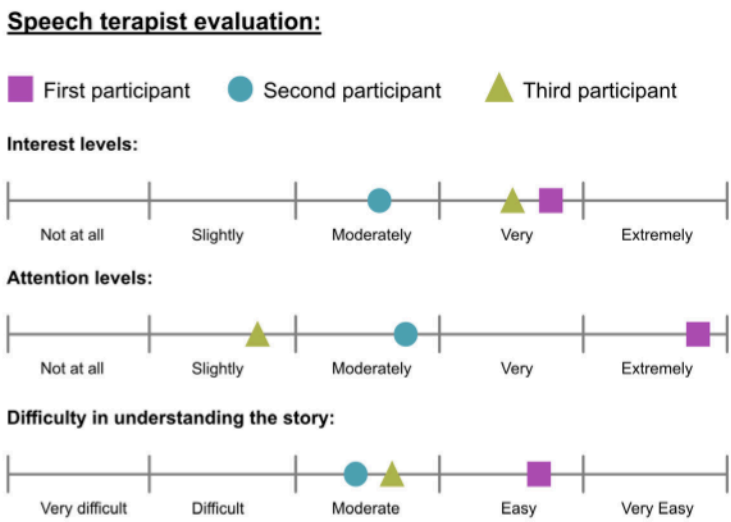


Fig. 4
Speech-language therapists' ratings of the three participants across three dimensions: interest, attention, and difficulty in understanding the story.

diid No. 89 — 2026
Doi: 10.30682/diid8926i

The embodied conversational agent was displayed on a computer screen and operated by the therapist using a mouse or keyboard. During initial interactions, participants attempted to touch the screen directly, treating it as a touchscreen device, reflecting limited familiarity with desktop computer interaction and more exposure with touch screen displays. This observation suggests the need for further design consideration regarding input modalities and interaction affordances in future iterations of the system.

As shown in **Fig. 5**, interest and attention did not evolve uniformly across the three participants. The first participant showed high interest and sustained attention, whereas the second increased attention after the beginning of the narrative. The third participant initially showed high interest but progressively disengaged. These differences reinforce the need to consider narrative engagement as contingent on individual interactional conditions rather than as a universally effective motivational strategy.

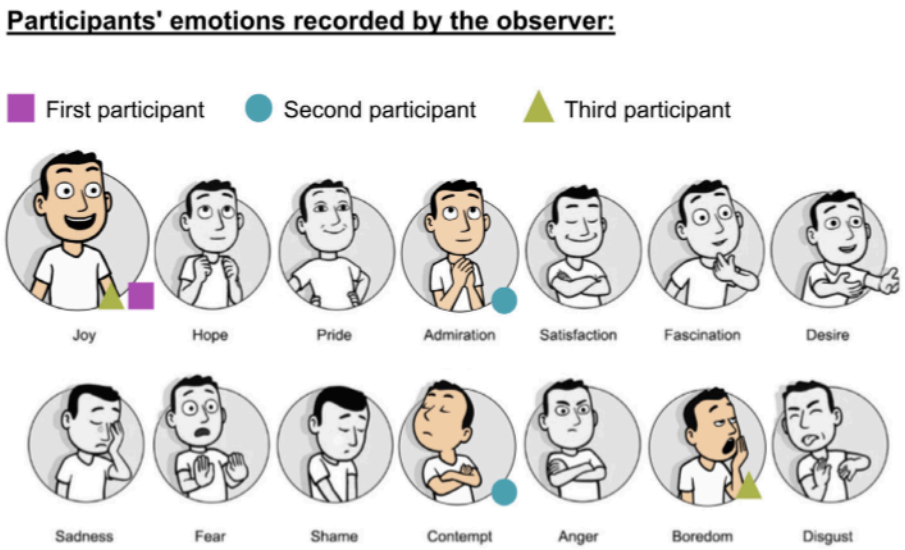


Fig. 5
Emotional responses (PrEmo2) recorded in the pilot show variable distribution across the three interactional profiles.

Discussion

These findings of the study were translated into six main design principles: non-judgmental error mediation, adaptive scaffolding, narrative reframing, and calibrated relational embodiment, flexible engagement and autonomous home practice. The first principle informed Pierina's conversational behaviour: rather than explicitly evaluating errors, the agent provides encouragement, suggestions, and opportunities to try again. The second translated the therapist's adaptive practice into mechanisms for modulating task difficulty, allowing the system to return to simpler activities when persistent difficulties occur. The third principle informed the integration of speech exercises within the narrative of Ziza's Words, transforming repetitive therapeutic tasks into meaningful actions within the game world.

The fifth principle suggested that storytelling should be flexible and calibrated to children's individual abilities and interaction preferences rather than assumed to be universally motivating. The last one focused on autonomous practice at home, to allow therapists to propose exercises to be carried out autonomously, and to monitor the activities performed outside the clinical setting, thereby offering a comprehensive overview of the child's rehabilitation pathway.

Calibrated relational embodiment represents a further design insight emerging from the study. Pierina was deliberately conceived as a playful companion rather than an authority figure or evaluator. Embodiment is therefore treated not simply as a means of increasing social presence, but as a design variable through which the appropriate relational distance between child and agent can be established. This approach responds also to ethical concerns concerning excessive emotional attachment and dependency in interactions with embodied agents.

The prototype consequently functions not only as a technological outcome but as a research artefact through which empirical observations are translated into design knowledge. The contribution of the study lies in these preliminary design propositions for emotionally sensitive ECAs, rather than in demonstrating clinical effectiveness. The formative pilot, involving only three children and focusing on the /s/ phoneme, provided initial insights into interest, attention, emotional responses, and interaction affordances, but cannot support claims about therapeutic efficacy or generalisation. The unexpected tendency of participants to touch the computer screen also revealed the need to reconsider input modalities in future iterations.

Overall, the study suggests that the design of conversational technologies for paediatric rehabilitation should address not only the delivery of therapeutic exercises, but also the emotional and relational conditions surrounding them. The proposed principles provide a preliminary basis for further research on ECAs that complement, rather than replace, the therapist's role, while future studies with larger and more diverse samples will be required to assess their applicability and potential therapeutic effects.

Conclusion

This paper presented a design-driven investigation into the development of a gamified speech therapy system supported by an embodied conversational agent for paediatric patients. Through a qualitative multi-method approach combining literature review, field observations of clinical speech therapy sessions and interviews with therapists, the study identified key experiential and relational challenges associated with conventional therapeutic practices, particularly in home-based contexts. These insights informed the development of a game integrating narrative, feedback mechanisms, and embodied interaction, while the design process articulated how therapeutic knowledge could be translated into the agent's embodiment, interaction style, and coaching role.

Patrizia Marti
She is a Professor of Experience Design at the Università degli Studi di Siena. She is the Rector's delegate for the Santa Chiara Fab Lab (www.scfablab.unisi.it) where she manages participatory innovation projects. She is a member of the Board of Directors of SID – Società Italiana di Design. From 2013 to 2019 she was appointed Professor in the Department of Industrial Design, Eindhoven University of Technology.
She has an interdisciplinary background in design, philosophy and computing. Her research focuses on designing technological systems in sectors such as health, accessibility, cultural heritage, and food. She studies the impact of digital technologies on humans, society and the environment.

Alessia Nicoletta Marino
PhD student in the Italian National Doctoral Programme Design for Made in Italy: Identity, Innovation, and Sustainability at the Università degli Studi di Siena and the Università degli Studi della Campania "Luigi Vanvitelli". Her academic path has led her to explore the fields of User Experience Design, Interface Design, and Conversational Design. Her research focuses particularly on AI-based virtual agents.

Andrea Brogi
He is a Speech and Language Therapist and Clinical Psychologist. He is currently training as a psychotherapist at the SPC School of Psychotherapy. He is a member of AIRIPA (Italian Association for Research and Intervention in the Psychopathology of Learning). He has several years of experience in private clinical practice, where he focuses on the assessment and treatment of neurodevelopmental disorders, with particular attention to Specific Learning Disorders and primary language disorders. His clinical and professional interests centre on evidence-based assessment and intervention for developmental and learning-related difficulties.

diid No. 89 — 2026
Doi: 10.30682/diid8926i

The study shows how design research can make visible and translate into technological artefacts the relational and emotional knowledge embedded in paediatric speech therapy practice. Rather than proposing the embodied conversational agent as a substitute for the therapist, the study explores how selected aspects of therapeutic interaction—particularly the management of errors, adaptation, encouragement, and relational positioning—can be incorporated into the design of home-based therapeutic experiences.

The resulting prototype should be considered a design research artefact and a set of preliminary design propositions, rather than a validated therapeutic intervention. The exploratory pilot provided initial indications of how children may engage differently with narrative, embodiment, and interaction, while also revealing aspects requiring further design iteration. Future research will investigate these propositions with larger and more diverse groups of children and through evaluations capable of addressing both interactional and therapeutic outcomes.

The study ultimately suggests that the contribution of design in digital rehabilitation lies not only in making therapy more engaging, but in designing the quality of the relationships that technology mediates.

Reference

Alsebayel, G., Nasri, M., Myers, C. P., Troiano, G. M., Hatamimajoumerd, E., Ostadabbas, S., Allison, K., & Harteveld, C. (2024). Articumotion: Towards assessing motor speech disorders via gamification. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference* (pp. 232-247). https://doi.org/10.1145/3628516.3655581

Antunes, I., Antunes, A., & Madeira, R. N. (2023). Designing serious games with a pervasive therapist interface for phonetic-phonological assessment of children. In *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia* (pp. 532–534). https://doi.org/10.1145/3626705.3631798

Busch, I. M., Moretti, F., Travaini, G., Wu, A. W., & Rimondini, M. (2019). Humanization of care: Key elements identified by patients, caregivers, and healthcare providers. A systematic review. *The Patient-Patient-Centered Outcomes Research*, *12*(5), 461–474. https://doi.org/10.1007/s40271-019-00375-6

Carolus, A., & Wienrich, C. (2022). "Imagine this smart speaker to have a body": An analysis of the external appearances and the characteristics that people associate with voice assistants. *Frontiers in Computer Science*, *4*, 981435. https://doi.org/10.3389/fcomp.2022.981435

Cassell, J. (2001). Embodied conversational agents: representation and intelligence in user interfaces. *AI magazine*, *22*(4), 67-83. https://doi.org/10.1609/aimag.v22i4.1593

Dalsgaard, P. (2010). Research in and through design: an interaction design research approach. In *Proceedings of the 22nd conference of the Computer-Human Interaction special interest group of Australia on Computer-Human Interaction* (pp. 200-203). https://doi.org/10.1145/1952222.1952265

Desmet, P. (2018). Measuring emotion: Development and application of an instrument to measure emotional responses to products. In M. Blythe, & A. Monk (Eds.), *Funology 2: From usability to enjoyment* (pp. 391–404). Springer International Publishing.

Forlizzi, J., Zimmerman, J., Mancuso, V., & Kwak, S. (2007). How interface agents affect interaction between humans and computers. In *Proceedings of the 2007 conference on Designing pleasurable products and interfaces* (pp. 209–221). https://doi.org/10.1145/1314161.1314197

Garg, R., Cui, H., Seligson, S., Zhang, B., Porcheron, M., Clark, L., Cowan, B. R., & Beneteau, E. (2022). The last decade of HCI research on children and voice-based conversational agents. In *Proceedings of the 2022 CHI conference on human factors in computing systems* (pp. 1–19). https://doi.org/10.1145/3491102.3502016

Hair, A., Monroe, P., Ahmed, B., Ballard, K. J., & Gutierrez-Osuna, R. (2018). Apraxia world: A speech therapy game for children with speech sound disorders. In *Proceedings of the 17th ACM Conference on Interaction Design and Children* (pp. 119-131). https://doi.org/10.1145/3202185.3202273

Imaezue, G. C., & Marampelly, H. (2025). ABCD: A simulation method for accelerating conversational agents with applications in aphasia therapy. *Journal of Speech, Language, and Hearing Research, 68*(7), 3322-3336. https://doi.org/10.1044/2025_JSLHR-25-0000

Khan, F., Wu, Y., Dray, J., Hemsley, B., & Kocaballi, A. B. (2025). Conversational agents to support people with communication disability: A co-design study with speech pathologists. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1-9). https://doi.org/10.1145/3706599.3719800

Koskinen, I., Zimmerman, J., Binder, T., Redström, J., & Wensveen, S. (2011). *Design research through practice: From the lab, field, and showroom*. Morgan Kaufmann.

Kováč, J., & Ondáš, S. (2025). LogoHra: A customizable gamified speech therapy platform. In *2025 International Conference on Emerging eLearning Technologies and Applications* (ICETA) (pp. 456-461). IEEE. https://doi.org/10.1109/ICETA67772.2025.11280071

Lan, T., Aryal, S., Ahmed, B., Ballard, K., & Gutierrez-Osuna, R. (2014). Flappy Voice: an interactive game for childhood apraxia of speech therapy. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play* (pp. 429-430). https://doi.org/10.1145/2658537.2661305

Murtarelli, G., Gregory, A., & Romenti, S. (2021). A conversation-based perspective for shaping ethical human–machine interactions: The particular challenge of chatbots. *Journal of Business Research*, *129*, 927–935. https://doi.org/10.1016/j.jbusres.2020.09.018

Oliveira, J., Silva, T., Oliveira, R., & Furtado, E. (2025). Recommendations of embodied conversational agents to healthcare applications. In *Proceedings of the 11th International Conference on Information and Communication Technologies for Ageing Well and e-Health* (ICT4AWE 2025) (pp. 352–359). SCITEPRESS. https://doi.org/10.5220/0013438700003938

Perrotta, E., & Rustici, P. (2023). *Correggere i difetti di pronuncia. Il programma A.P.I. (Ascolta-Prova-Impara) per l'allenamento percettivo-articolatorio* [Correcting pronunciation defects. The A.P.I. program (Listen-Try-Learn) for perceptual-articulatory training]. Erickson.

Poushneh, A. (2021). Humanizing voice assistant: The impact of voice assistant personality on consumers' attitudes and behaviors. *Journal of Retailing and Consumer Services*, *58*. https://doi.org/10.1016/j.jretconser.2020.102283

Qiu, L., & Abdullah, S. (2021). Voice assistants for speech therapy. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers* (pp. 211-214). https://doi.org/10.1145/3460418.3479336

Read, J. C., & Markopoulos, P. (2013). Child–computer interaction. *International Journal of Child-Computer Interaction, 1*(1), 2–6. https://doi.org/10.1016/j.ijcci.2012.09.001

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places.* Cambridge University Press.

Rinaldi, S., Caselli, M. C., Cofelice, V., D'Amico, S., De Cagno, A. G., Della Corte, G.,Di Martino, M. V., Di Costanzo, B., Levorato, M. C., Penge, R., Rossetto, T., Sansavini, A., Vecchi, S., & Zoccolotti, P. (2021). Efficacy of the treatment of developmental language disorder: A systematic review. *Brain Sciences*, *11*(3). https://doi.org/10.3390/brainsci11030407

Saeedi, S., Bouraghi, H., Seifpanahi, M. S., & Ghazisaeedi, M. (2022). Application of digital games for speech therapy in children: A systematic review of features and challenges. *Journal of Healthcare Engineering, 2022.* https://doi.org/10.1155/2022/4814945

Spitale, M., Silleresi, S., Cosentino, G., Panzeri, F., & Garzotto, F. (2020, June). "Whom would you like to talk with?": exploring conversational agents for children's linguistic assessment. In *Proceedings of the interaction design and children conference* (pp. 262-272). https://doi.org/10.1145/3392063.3394422

Teverini, G., Recupero, A., & Marti, P. (2025). Therapy gamification to improve children compliance: A systematic literature review. *ID&A Interaction Design & Architecture(S), 65*, 257–281. https://doi.org/10.55612/s-5002-065-009

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 493-502). https://doi.org/10.1145/1240624.1240704

diid No. 89 — 2026
Doi: 10.30682/diid8926i