

SentenceBot: An Adaptive Social Robot for Language Therapy in Children with Developmental Language Disorder

Individual Portfolio

Name	Baran Akkanat
Student Number	
Group	Group 3
Course	HRI for Social Robots
Lecturer	Paul Vogt
Institution	University of Groningen
Deadline	17 June 2026

Brief Summary

My contribution to SentenceBot centred on three interconnected components: the adaptive difficulty engine, the QR card topic selection system, and the celebration gesture implementation. I also acted as the child user in the video demonstration, which constitutes my primary piece of individual evidence.

Adaptive Difficulty Engine

The adaptive difficulty engine forms the core therapeutic logic of SentenceBot. I designed and implemented the three-stage progression system in which vocabulary is partitioned into easy words (high-frequency, phonologically simple items such as *dog* and *sun*) and hard words (lower-frequency items such as *elephant* and *thunder*), with a third stage requiring the child to incorporate two randomly selected words into a single coherent sentence. The session comprises three rounds of two questions each; advancing to the next stage requires both questions in a round to be answered correctly, while an incorrect answer holds the difficulty at the current level.

This mastery-based criterion was grounded in Vygotsky’s zone of proximal development [1]: tasks should be calibrated to remain slightly beyond current performance while still being achievable, so that the interaction provides genuine cognitive challenge without inducing frustration. I also implemented the streak detection logic, which triggers a special encouragement message after every third consecutive correct answer, consistent with variable reward schedules shown to sustain engagement in child-robot interaction [2].

QR Card Topic Selection

The decision to use physical QR cards rather than a spoken or touch-based menu was motivated directly by Azizi et al. [3], who showed through co-design with speech-language pathologists that tangible, physical materials reduce cognitive load for children with communication difficulties. I implemented the OpenCV-based decoding pipeline that reads the robot’s front-facing camera feed via the `rom.sensor.sight.read` API and decodes the topic name using `QRCodeDetector`. I also designed and produced the four printable QR cards in $\text{\LaTeX}$ using the `qrcode` package — one card per topic (animals, food, school, weather) — and laminated them for repeated use across testing sessions. Detection proved reliable at 20–60 cm under normal indoor lighting conditions throughout all testing sessions.

Celebration Gesture

The celebration gesture was implemented using direct motor control via the `rom.actuator.motor.write` API. On receipt of a positive validity flag from the LLM validator, the robot raises both arms simultaneously and then returns to the neutral standing pose.

The choice of a bilateral arm raise was made after considering three alternatives: a single-arm wave, a head nod, and a torso lean forward. The single-arm wave was rejected because it risked being interpreted as a greeting rather than an expression of praise, which could confuse the child about the turn structure. Head nod and torso lean were excluded because their amplitude on the Alpha Mini is too small to be salient from the child’s seated position. Raising both arms produces the largest and most symmetrical motion the platform supports and maps intuitively onto the culturally familiar “hands up” celebration gesture. This multimodal response — combining LLM-generated verbal encouragement with a physically salient gesture — reflects evidence that gesture-accompanied feedback enhances engagement and perceived warmth in child-robot interaction [2].

Video Demonstration

I participated as the child user in the video demonstration, responding verbally to the robot’s prompts across all three adaptive difficulty rounds. Adopting the child role rather than operating the robot from behind the camera was a deliberate methodological choice: it forced me to experience the interaction from the perspective of the end user and, in doing so, surfaced a latency issue in the LLM validation loop that had not been apparent during developer-side testing. The delay between the child’s spoken answer and the robot’s verbal response was long enough to feel unnatural in real interaction, which led the group to add a brief filler utterance (“Let me think...”) to bridge the gap. This adjustment would not have been identified had the video been recorded with a non-group member acting as the child. The recording captures QR card topic selection, the complete difficulty progression, the STT pipeline under natural speech conditions, and the known misrecognition limitation (e.g., *rice* transcribed as *price*) in a single continuous take, and it is submitted alongside this portfolio as my primary individual evidence.

Individual Reflection

What I Learned

This project gave me my first sustained exposure to event-driven, asynchronous Python programming through the WAMP/Twisted framework. Understanding how the robot’s perception, cognition, and action layers communicate through asynchronous message passing required a significantly different mental model from the synchronous scripts I had

written before. The first week of the implementation phase was largely consumed by this learning curve, and working through it was one of the most technically demanding experiences of the course.

Beyond the technical side, engaging with the clinical DLD literature deepened my understanding of what it actually means to design for a clinical population. Reading Bishop et al. [4] and Rice and Wexler [5] changed how I thought about validation: accepting morphologically incomplete but communicatively valid sentences as correct is not a design shortcut but a clinically motivated decision, because children with DLD characteristically omit tense morphology well beyond the age at which typically developing children resolve this. That connection between the literature and a concrete implementation detail was one of the more satisfying intellectual experiences of the project.

How I Met the Learning Goals

The course learning goals include understanding HRI design principles, applying challenge-based learning through iterative design, and reflecting critically on both the process and the product.

On HRI design principles: the adaptive engine instantiates Vygotsky's zone of proximal development as a concrete branching algorithm, and the QR card interface embeds the finding from Azizi et al. that tangible materials lower cognitive load for children with communication difficulties. These were not post-hoc justifications; both decisions were made before implementation began, with the literature guiding the design choice.

On challenge-based learning through iterative design: the STT misrecognition problem was identified mid-project during live testing, not anticipated in the initial design. Documenting it honestly in the video rather than re-recording a clean run reflects the iterative spirit the course promotes — understanding the system's failure modes is part of the design process.

On critical reflection: the limitations section below identifies the most consequential gap (absence of clinical evaluation) and proposes a concrete next step, rather than treating the current prototype as a finished product.

Role in the Group

My role was primarily in the implementation of the game logic and the physical interface components. One example of how the group resolved disagreement constructively: there was an early debate about whether the difficulty progression should be time-based (advancing after a fixed number of seconds regardless of correctness) or mastery-based (advancing only after both questions in a round are answered correctly). I ran a small informal test with the two approaches using placeholder speech output and showed the

group that a time-based system produced frustrating transitions mid-attempt; this resolved the debate in favour of the mastery criterion without further discussion. By the end of the project all three members understood the complete codebase, which made integration testing significantly more efficient.

Limitations and Future Directions

The most significant limitation I encountered directly was the STT misrecognition problem, which surfaced during testing and which I documented in the video evidence. The Alpha Mini’s built-in microphone was not designed for child speech and struggled with phonologically similar word pairs. Future work should explore ASR models specifically trained on child speech corpora [6], which would substantially improve the system’s reliability in a clinical setting. The absence of formal evaluation with DLD-diagnosed children also limits the evidence base for our therapeutic claims, and I see clinical validation as the most important next step.

Evidence

- **Video demonstration** (submitted with portfolio) — Full interaction recording capturing QR card topic selection, all three adaptive difficulty rounds, LLM validation feedback, celebration gesture, and STT misrecognition example. I acted as the child user throughout the recording.
- `main.py` (submitted in source code zip) — Contains the adaptive difficulty engine (stages 1–3, mastery-based advancement, streak detection), QR decoding pipeline (OpenCV `QRCodeDetector` + `rom.sensor.sight.read`), and celebration gesture (`rom.actuator.m`).
- `qr_cards.tex` (submitted in source code zip) — $\text{\LaTeX}$ source for the four QR topic cards (animals, food, school, weather), produced using the `qrcode` package and printed and laminated for physical use during testing.

References

- [1] L. S. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press, 1978.
- [2] T. Belpaeme et al., “Social robots for education: A review,” *Science Robotics*, vol. 3, no. 21, p. eaat5954, 2018.
- [3] N. Azizi et al., “Development of robot-assisted speech-language therapy: Co-design with speech-language pathologists,” in *Proc. ICSR*, Springer, pp. 366–380, 2025.
- [4] D. V. M. Bishop et al., “Phase 1 of CATALISE: A multinational Delphi consensus study identifying language impairments in children,” *PLOS ONE*, vol. 11, no. 7, p. e0158753, 2016.
- [5] M. L. Rice and K. Wexler, “Toward tense as a clinical marker of specific language impairment,” *Journal of Speech, Language, and Hearing Research*, vol. 39, no. 6, pp. 1239–1257, 1996.
- [6] M. Spitale et al., “Using socially assistive robots in speech-language therapy for children with language impairments,” *International Journal of Social Robotics*, vol. 15, no. 9, pp. 1525–1542, 2023.