POLITECNICO MILANO 1863

Designing a physical interaction layer for a multimodal Storytelling App for children users of AAC

Master thesis - AA 2025/2026

Digital Interaction Design Master Degree

Politecnico di Milano

Nina Alejandra Mesias Barrera | 11061871

Supervisor: Prof. Franca Garzotto

Co-Supervisors: Prof. Daniela Petrelli,

Giovanni Caleffi, Marta Curreri, Giulia Valcamonica

TABLE OF CONTENTS

LIST OF FIGURES 4

ABSTRACT EN 8

ABSTRACT IT 9

1. INTRODUCTION 11

2. RESEARCH FRAMING 13

    2.1 Augmentative and Alternative Communication 13

    2.2 Minimally verbal children 15

    2.3 Assistive Technology 16

    2.4 Storytelling in child development and AAC therapy 19

    2.5 Tangible User Interfaces 20

3. LITERATURE REVIEW 24

  3.1. CASE STUDY ANALYSIS 24

    Personalized storybooks using generative AI: Autihero 24

    "NarrAid: Supporting Storytelling of People with Aphasia via Generative Visual Scene Displays"[22] 25

    "Moovy: a Tangible User Interface to Train Linguistic Skills in Children"[23] 27

    Linguabytes: Tangibles for toddlers learning language 28

  3.2. BENCHMARKING 30

    3.2.1. AAC APPS 30

    AVAZ 31

    Proloquo2Go 32

4. PROJECT DEFINITION 36

    4.1. AACREATIVE: THE STARTING POINT 36

    4.2. UNDERSTANDING THE CONTEXT OF USE 41

4.2.1. Exploratory Digital Content Review 41
4.2.2. Interviews 44
4.2.3. Shared therapy sessions 49
5. DESIGN PROCESS 51
5.1. VERSION 1 51
5.2. VERSION 2 56
5.3 CODESING WORKSHOP 65
5.4. INSIGHTS 68
6. DESIGN ADJUSTMENTS PROPOSALS 71
6.1. SUGGESTIONS FOR REDESIGNING THE AACREATIVE USER EXPERIENCE 71
6.2. ADJUSTMENTS TO THE PHYSICAL SETUP 77
6.3. ADDITIONAL PROPOSALS 79
7. LIMITATIONS AND FUTURE WORK 80
8. CONCLUSIONS 84
BIBLIOGRAPHY 86

LIST OF FIGURES

Figure 1. "The four components of the Human Activity Assistive Technology (HAAT) model"[2]....17

Figure 2. “Alternative access to computing devices. Note user controls an access technology interface to use a computing device for desired activities.”. [9] 18

Figure 3a,b. GUI and TUI interaction models. [16]....21

Figure 4. a. (left) GUI-based interfaces lean heavily on the child's cognitive skills. b.(right)TUIs open up skills through perceptual-motor skills. [17] 22

Figure 5. User Interface of Autihero. In this section the caregiver gives details about the child. [19]....25

Figure 6. “Interface of NarrAid. (A) AI-generated guidance. Here (a1) is Mode-switching bar between free-chat mode and guidance mode. (B) The Input module includes: (b1) input bar with delete button, (b2) image card selection (pictographic symbols are sourced from ARASAAC), (b3) core buttons including Yes, No, Skip, Not clear, and Generate, and (b4) custom image card component allowing caregivers to upload personalized cards which then appear in (b2). (C) The window to display high-context images generated by AI. (D) My Stories archive to store generated images, texts, and cards in a card-style layout” [22] 27

Figure 7. User Interface of AVAZ, App for communicating using pictograms.[26]....32

Figure 8. Proloquo2Go User Interface and its functionalities [30]. 34

Figure 9. Other Apps explored: Left to right, top to bottom: Commboards AAC, Otsimo AAC, Larunda and DictaPicto. 35

Figure 10. AACreative app: Caregiver section....37

Figure 11. AACreative App: Creation of a personalized card....37

Figure 12. Overlay with the sets created by the caregiver. 38

Figure 13. AACreative App: Create section, designed for the AAC user....39

Figure 14. AACreative: Screen for waiting the generation of the story. 39

Figure 15. AACreative: Going through the multimodal story 40

Figure 16. AACreative: The Library section....40

Figure 17. Example of AAC used in different settings. Source: Youtube....43

Figure 18. Examples of use and preparation of AAC pictograms using low tech aids. Source: Pinterest....43

Figure 19. Remote Interview with a speech-language pathologist....48

Figure 20. Example of dices used for story generation in the context of speech therapy. 50

Figure 21. Ideation process to generate a first design concept (right image generated with Gemini AI)....51

Figure 22. Hardware involved in physical sketch 1. 52

Figure 23. Screen for the caregiver, to associate the physical cards with their digital representation. 53

Figure 24. Screen to select the digital representation of the last NFC tag read. Overlay of the caregiver section....53

Figure 25. Screen for the interaction of the AAC user....54

Figure 26. First design concept....55

Figure 27. Controller of the first design concept....55

Figure 28. Controller of prototype 2. Since in this version the ESP32 is not connected to the tablet a battery is needed. 57

Figure 29. Caregiver section where the "Cards with tag" filter appears to create an association with physical cards 58

Figure 30. Overlay for waiting the scanning process of a physical card. 59

Figure 31. Overlay to associate the physical card with its digital version. 59

Figure 32. The Physical Set appears as the first option in the Sets list if there is a Bluetooth connection stablished. 60

Figure 33. In the Create section, when a physical card is scanned it appears on the screen and is read aloud. 61

Figure 34. Once the "Create story" button has been scanned the creation process starts and an animation is shown, ...62

Figure 35. Story navigation can be done scanning the physical buttons....63

Figure 36. Communication between the AACreative App and the tangible controller. Image generated by Gemini AI based on a drawing by the author....64

Figure 37. Physical-Digital Interaction Map for the caregiver 64

Figure 38. Physical-Digital Interaction Map for the AAC user....65

Figure 39. Example of the use of personalized cards in the storybook 66

Figure 40. Participation in the workshop: Preparation of the cards for the story creation 68

Figure 41. Addition of the button "Associate tag" in the section of personalized cards. (Made with Gemini AI) 72

Figure 42. Colors used to categorize the words in the modified Fitzgerald keys system.[28] 72

Figure 43. Adjustment in the create section: Substitute the text button for ARASAAC pictograms. 73

Figure 44. First page of the story presents the title and the authors. 74

Figure 45. Proposal for rearrangement of the elements in the storybook (made with Gemini AI) 75

Figure 46. Page to close the story. 75

Figure 47. Proposal for the frames that would contain the pictograms. 78

Figure 48. Different ways of gripping that support motor impairment. Source: Pinterest 79

Figure 49. AAC dices to create random stories first by the AAC user, then by the system. (image generate by Gemini AI) 80

ABSTRACT EN

This thesis presents the design of a tangible interaction layer for AACreative, an existing tablet application that uses symbol-based input and AI-generated multimodal storytelling to support minimally verbal children who use Augmentative and Alternative Communication (AAC) systems. While AACreative successfully combines AAC pictograms with narrative generation, its reliance on touchscreen interaction limits accessibility for children with motor impairments and constrains the spatial, participatory dimension of the storytelling experience. This work investigates how a physical interaction layer, grounded in the theoretical tradition of Tangible User Interfaces, can extend the system's expressive and therapeutic potential.

The design process followed a user-centered, iterative approach structured around three interrelated profiles: the AAC user, the speech-language therapist, and the caregiver. Research methods included an exploratory review of digital content, semi-structured interviews with speech-language therapists, and a caregiver in Italy and Colombia, and the observation of recorded therapy sessions at the Istituto Neurologico Carlo Besta. These methods surfaced recurring themes, including the central role of the caregiver as mediator, the necessity of personalization at both visual and linguistic levels, the relevance of motor accessibility as a design constraint, and the therapeutic value of narrative as a higher-order communicative function.

Two sketches in hardware were developed to translate these findings into a working interaction model based on NFC-tagged physical cards read through an ESP32 microcontroller. The first established the core interaction flow using serial communication; the second integrated the physical layer directly into the AACreative production application via Bluetooth Low Energy, enabling the AAC user profile compose and navigate AI-generated stories entirely through tangible manipulation. A co-creation session held at the Istituto Neurologico Carlo Besta brought the second sketch into direct contact with speech-language therapists, generating insights that informed a series of proposed adjustments to both the physical setup and the AACreative application's user experience, including card categorization aligned with narrative composition questions, color-coding for grammatical roles, restructured storybook pages, and refinements to the story generation prompt aimed at improving narrative coherence.

ABSTRACT IT

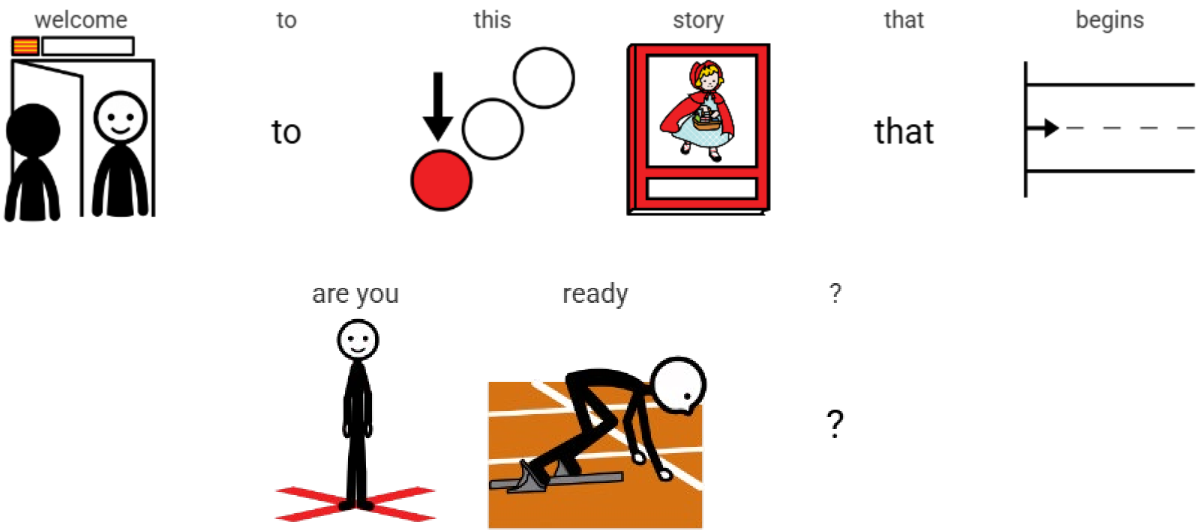
Questa tesi presenta la progettazione di uno strato di interazione tangibile per AACreative, un'applicazione per tablet già esistente che utilizza input basati su simboli e narrazione multimodale generata tramite intelligenza artificiale per supportare bambini minimamente verbali che utilizzano sistemi di Comunicazione Aumentativa e Alternativa (CAA). Sebbene AACreative combini efficacemente i pittogrammi CAA con la generazione narrativa, la sua dipendenza dall'interazione tramite touchscreen limita l'accessibilità per i bambini con compromissioni motorie e riduce la dimensione spaziale e partecipativa dell'esperienza narrativa. Questo lavoro indaga come uno strato di interazione fisica, fondato sulla tradizione teorica delle Tangible User Interfaces, possa estendere il potenziale espressivo e terapeutico del sistema.

Il processo di progettazione ha seguito un approccio iterativo e centrato sull'utente, strutturato attorno a tre profili interconnessi: l'utente CAA, il logopedista e il caregiver. I metodi di ricerca includono una revisione esplorativa di contenuti digitali, interviste semi-strutturate con logopediste e un caregiver in Italia e in Colombia, e l'osservazione di sessioni terapeutiche registrate presso l'Istituto Neurologico Carlo Besta. Questi metodi hanno fatto emergere temi ricorrenti, tra cui il ruolo centrale del caregiver come mediatore, la necessità di personalizzazione sia a livello visivo che linguistico, la rilevanza dell'accessibilità motoria come vincolo progettuale, e il valore terapeutico della narrazione come funzione comunicativa di livello superiore.

Sono stati sviluppati due sketch in hardware per tradurre questi risultati in un modello di interazione funzionante, basato su carte fisiche dotate di tag NFC lette tramite un microcontrollore ESP32. Il primo ha definito il flusso di interazione di base utilizzando comunicazione seriale; il secondo ha integrato lo strato fisico direttamente nell'applicazione AACreative in produzione tramite Bluetooth Low Energy, consentendo al profilo dell'utente CAA di comporre e navigare storie generate dall'intelligenza artificiale interamente attraverso la manipolazione tangibile. Una sessione di co-creazione svolta presso l'Istituto Neurologico Carlo Besta ha messo il secondo sketch a diretto contatto con logopediste, generando spunti che hanno orientato una serie di proposte di adeguamento sia al setup fisico sia all'esperienza utente dell'applicazione AACreative, tra cui la categorizzazione delle carte in linea con le domande di composizione narrativa, una

codifica a colori per i ruoli grammaticali, una riorganizzazione delle pagine del libro di storie, e perfezionamenti del prompt di generazione narrativa volti a migliorare la coerenza della storia.

1. INTRODUCTION



Minimally verbal children who rely on Augmentative and Alternative Communication (AAC) systems are increasingly supported by digital tools that allow them to communicate using symbol-based input and speech generation as output. Yet most of these tools remain confined to the touchscreen, a mode of interaction that can be challenging for children with motor impairments, and that offers a limited range of expressive possibilities. At the same time, storytelling is widely recognized in clinical practice as a higher-order communicative function, one that supports vocabulary generalization, sequencing, and shared attention between child and caregiver, yet it remains underexplored as a structured feature of AAC technology. This thesis investigates how a tangible interaction layer might address both of these gaps, taking as its case study AACreative, an existing application that generates multimodal illustrated stories from AAC symbol input.

AACreative is a tablet application that uses generative artificial intelligence and ARASAAC [1] (Aragonese Center of Augmentative and Alternative Communication) pictograms to produce multimodal illustrated stories for minimally verbal children who use AAC systems. It was originally developed as a final project for the Advanced User Interfaces course at Politecnico di Milano by a team of four students and represents a complete system with a caregiver configuration interface, a creation section managed by the AAC user, and a story library. This thesis enters the project at a subsequent stage, proposing an extension to the existing system rather than designing it from scratch.

The central proposition of this work is that adding a physical interaction layer to AACreative can meaningfully expand its expressive and therapeutic potential. This proposal draws on the body of research around Tangible User Interfaces (TUI), which this thesis examines to apply it to the specific context of AAC and minimally verbal children. For children who also present motor or attentional challenges, the possibility of interacting with a digital system through physical objects rather than a touchscreen raises questions that go beyond usability, touching on the qualitative nature of the interaction itself.

The design process followed in this thesis is user-centered and iterative, structured around three distinct user profiles with different needs: the AAC user, who is the primary beneficiary of the system; the therapist or speech-language pathologist, who defines the therapeutic objectives and mediates the interaction; and the caregiver, who can create personalized cards in the system and accompanies the child during use. Understanding these three profiles and their relationships was the first phase of the process, carried out through an exploratory review of digital content, semi-structured interviews with professionals and one parent, and the observation of recorded therapy sessions.

This investigative phase informed the construction of two sketches in hardware, developed not as finished products to be tested against predefined criteria, but as communication tools designed to make the design proposal concrete enough to be discussed, questioned, and refined in dialogue with clinical experts. A co-creation session held at the Istituto Neurologico Carlo Besta brought the second sketch in hardware into direct contact with speech-language therapists, generating insights that shaped subsequent adjustments to both the physical design and the application interface.

The thesis is organized as follows. Chapter 2 presents the concepts that frame the project. Chapter 3 reviews relevant case studies and existing AAC applications. Chapter 4 introduces AACreative in its current form and presents the user research process and its findings. Chapter 5 documents the two sketches in hardware and the co-creation session that informed the design adjustments proposed in Chapter 6. The thesis closes with limitations, conclusions, and directions for future development.

2. RESEARCH FRAMING

2.1 Augmentative and Alternative Communication

The American Speech Language and Hearing Association (ASHA) defines Augmentative and Alternative Communication (AAC) as “a set of procedures and processes by which an individual’s communication skills (i.e., production as well as comprehension) can be maximized for functional and effective communication”, so AAC refers to “**forms of expression different from spoken language**, that aim at increasing (augmentative) and/or compensate (alternative) the difficulties of communication and language of many people with disabilities” [1]. The distinction between both terms is meaningful: augmentative strategies build upon residual oral communication capacities, while alternative strategies provide a complete substitute when speech is not functional. [1]

AAC intervention serves a diverse population that can be broadly organized into three groups according to the nature and duration of their communicative needs. The first group includes individuals who have developed a solid understanding of language but face significant difficulties in spoken expression, their cognitive grasp of communication is intact, but the channel through which they can convey it is limited (this is the case of people with apraxia of speech). The second group encompasses those who experience challenges both in comprehending speech and in producing it, often requiring structured support to build a symbolic or pictographic vocabulary that makes communication possible. A third group uses AAC on a temporary basis, typically children whose communicative development is still in progress or adults recovering from medical interventions that have temporarily affected their speech. Unlike the other two groups, these users may eventually transition away from AAC support as their primary communication abilities develop or are restored [2].

Communication and language are essential for every human being to interact with others, to learn, to enjoy, and to participate in society. When these capacities are impaired, the consequences extend well beyond language itself, affecting autonomy, social inclusion, and cognitive development. Among the **conditions** that may require the use of AAC are “cerebral palsy (CP), intellectual disability, autism spectrum disorders (ASD), neurological diseases such as amyotrophic lateral sclerosis (ALS), multiple sclerosis (MS) or Parkinson's disease, muscular dystrophies, traumatic brain injuries, aphasias or multiple disabilities of different types"[1]

Symbol Systems

AAC encompasses different systems of symbols, both graphic, such as pictures, drawings, pictograms, words, or letters; and gestural, such as mimicry, gestures, or hand signals. The different symbol systems adapt to the needs of people of varied ages and uneven motor, cognitive, and linguistic skills.

Within graphic systems, pictographic systems occupy a central role for children and individuals who have not yet developed language and communication skills or whose disability prevents their development. One of the key strengths of graphic symbol systems is their scalability across communicative levels. At one end of the spectrum, they can be configured for individuals with very limited cognitive development or those at the earliest stages of communication, offering a minimal but functional set of symbols. At the other end, they can support increasingly complex and nuanced expression as the user's competencies grow. That said, even in their most advanced configurations, graphic symbol systems do not fully replicate the expressive range and flexibility that written language affords [1]. Among the most widely used pictographic systems is **ARASAAC** (Aragonese Portal of Augmentative and Alternative Communication)[3], which provides a freely available symbol library under a Creative Commons license and constitutes the symbolic foundation of the AACreative application examined in this thesis [1].

Support Products and Access Strategies

Support products for AAC range from **non-technological** resources such as communication boards and books, to **technological** resources including electronic communicators, personal computers, and tablets with specialized software that enable different forms of access adapted to people with very limited mobility.

The modality of access to these systems is equally diverse. Strategies for accessing communication systems include **direct selection**, **mouse-based selection**, **dependent** or **independent scanning**, and **coded selection**[1]. AAC solutions are classified into no-tech, low-tech, and high-tech categories, with the latter encompassing a wide range of signal sensing approaches. These include imaging methods such as **eye gaze** and **head-pointing systems**, **touch-activated methods** via touchscreens and membrane keyboards, **mechanical and electromechanical switch** access, **breath-activated systems**, and **brain-computer interfaces (BCI)**[2]. Each modality presents different tradeoffs in terms of ease of access, affordability, portability, and communication speed, and the appropriateness of any given method is always contingent on the individual user's motor, cognitive, and contextual profile. [1], [2]

AAC as an Educational and Therapeutic Process

Beyond the technical dimension, the ARASAAC framework [3] emphasizes that tools alone are insufficient. Before any AAC system can be designed or recommended, a comprehensive assessment is necessary. This assessment must consider both the individual, their capacities, abilities, needs, and personal wishes, and the contexts in which they live and communicate, including the demands, resources, and constraints of their environment. Only from this dual understanding can the most appropriate system or combination of systems be determined [1]. This evaluation process should not be occasional but continuous throughout life. Furthermore, empowerment and teaching should be addressed both to the person and to their environment, across all the **contexts** where they participate, including professionals and family members, in educational and therapeutic settings as well as natural everyday contexts. [1]

This holistic perspective directly informs the design rationale of AACreative: the system was conceived not as an isolated technological artifact, but as a cooperative tool shared between **children and caregivers**, designed to be meaningful within the relational and educational contexts where communication development actually takes place.

2.2 Minimally verbal children

This term is generally reserved for children who reach **school age (age 5 and older)** and still have little to no expressive spoken language (fewer than 20 words). It is estimated that approximately **25–30% of children with Autism Spectrum Disorder (ASD)** fall into this category by the time they enter kindergarten.[4]

This population is notably heterogeneous; in part because this is a highly variable population with no single set of defining characteristics or patterns of skills or deficits, and in part because “it is extremely challenging to provide reliable or valid assessments of their developmental functioning” [4]. The label *minimally verbal* is therefore not a diagnostic category but a functional descriptor that captures a communicative profile shared across several conditions, including autism spectrum disorder, intellectual disability, and cerebral palsy. [4]

When children have not developed functional verbal communication by the age of five, the likelihood of lasting and significant communicative difficulties increases considerably. These difficulties rarely remain confined to language alone, they tend to affect the child's social relationships, their ability to participate and progress in educational settings, and their overall quality of life [5]. This trajectory makes early and sustained AAC intervention particularly urgent for this group, as the window for establishing alternative communication pathways is time-sensitive. [5]

For the purposes of this thesis, the term *minimally verbal* is used to describe the primary user population of AACreative: children who cannot rely on speech to communicate their needs, emotions, or creative expression, and who depend on symbol-based AAC systems as their primary expressive modality. The design of the physical device and the storytelling application is grounded in this communicative profile — prioritizing embodied, low-cognitive-load interaction that does not require literacy or verbal output to generate meaningful narrative.

2.3 Assistive Technology

Assistive technology (AT) refers to any product, device, software, or system that enables individuals with disabilities to perform functions that might otherwise be difficult or impossible, thereby supporting their participation in education, social life, and daily activities [6]. The World Health Organization recognizes that assistive products and technologies can enable people to live healthy, productive, independent and dignified lives, and to participate in education, the job market and civic life [7].

A key principle in AT design and selection is that the technology must be evaluated in relation to the **specific person,** activity, and context of use, a perspective captured in the Human Activity Assistive Technology (HAAT) model [2] (see Figure 1) which highlights that there is not general solutions since each one has to respond to someone doing something and using that specific assistive technology somewhere. Applied to children with complex communication needs, establishing an effective access method is a highly **individualized process** [8].

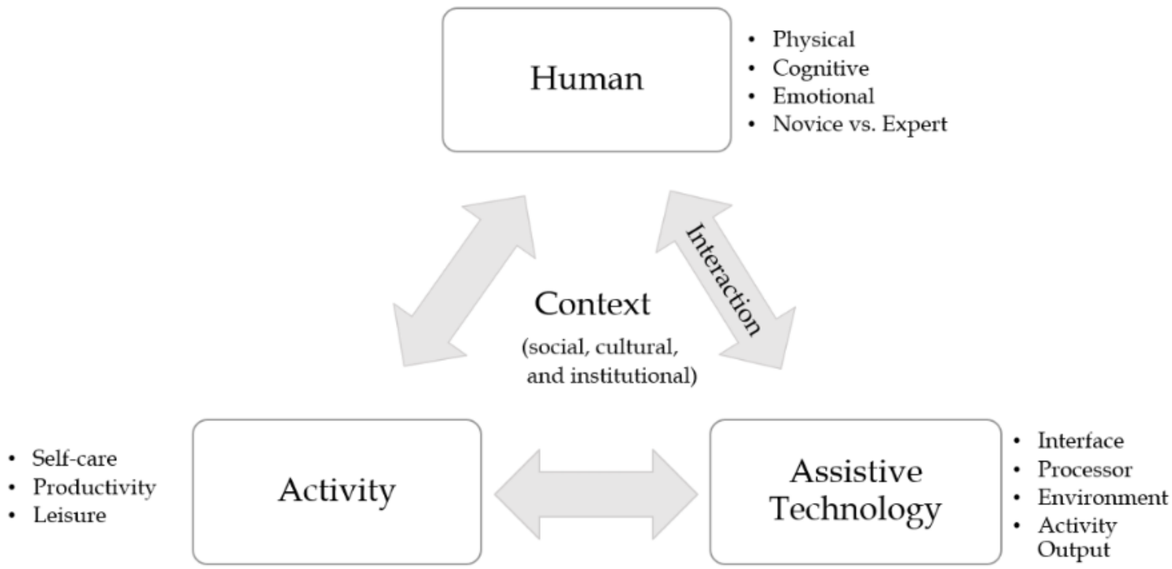


Figure 1. "The four components of the Human Activity Assistive Technology (HAAT) model"[2]

Within this broader AT landscape, AAC systems occupy a central position. Access methods fall under the broader category of assistive technologies [8], and their design directly shapes the communicative possibilities available to the child.

AAC Access Methods and the Limitations of Touchscreen Interaction

Within the AT landscape, AAC devices require not only appropriate symbol systems but also suitable **access methods** — the technologies and strategies that allow a user to interface with a device and make selections. As illustrated in Figure 2, the selection of an access method depends on the interplay between the user's available control sites (hand, head, eyes, speech), the control interface used, and the computing device being accessed.

Access methods are broadly classified as **direct** or **indirect** [8]. Direct selection involves pointing to or activating a target immediately, without navigating through other options;

typically achieved through touchscreens, keyboards, or eye-gaze systems. This method tends to offer the most efficient and intuitive interaction, with users potentially reaching selection speeds of 40–60 selections per minute that make real-time communication more viable[8]. **Indirect selection**, by contrast, requires the user to scan systematically through available options before making a choice, commonly through switch access. While indirect methods **reduce motor demands, they place greater load on attention, visual tracking, and working memory**, with typical rates of only 1–5 selections per minute.[8]

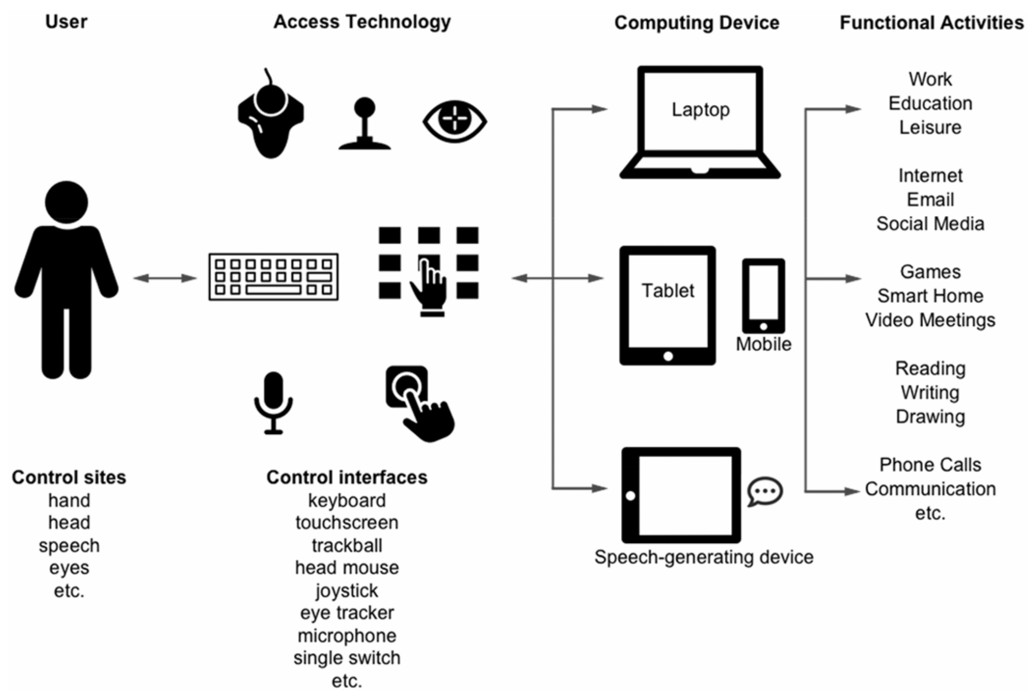


Figure 2. "Alternative access to computing devices. Note user controls an access technology interface to use a computing device for desired activities.". [9]

Many minimal verbal children have also motor impairment, that is the case of children with cerebral palsy, Angelman syndrome, Spinal muscular atrophy, dyspraxia among others, in this scenario standard touchscreen interfaces present significant challenges. Children with cerebral palsy frequently present motor control difficulties that directly affect their ability to use technology, including the very devices intended to compensate for their physical limitations. The challenges are multisensory and interconnected: disruptions in the processing of tactile, vestibular, and proprioceptive information, the persistence of primitive reflexes beyond their typical developmental window, atypical muscle tone, and the presence of involuntary movements or abnormal posturing all converge to complicate access to AAC systems in ways that cannot be addressed through software design alone. Even when touchscreen access is feasible, **the cognitive load** of navigating hierarchical symbol grids on a screen can overwhelm a child whose attentional and executive resources are already engaged in the communicative act itself.[8], [9]

This tension between **access demands and communicative intent** is precisely what motivates the tangible interaction approach explored in this thesis. Rather than requiring the child to point accurately at small targets on a screen, the NFC reader device proposed for AACreative shifts selection to the manipulation of distinct physical objects, a modality that leverages existing motor schemas, reduces visual scanning demands, and distributes the cognitive load of selection across the body and the environment.

2.4 Storytelling in child development and AAC therapy

Narrative, the ability to construct, sequence, and communicate a story, is far more than a literary skill. In child development, storytelling is recognized as a fundamental mechanism through which children develop communicative competence, build identity, process experience, and participate in social life. The ability to narrate a story is fundamental to the development of overall communicative competence and involves the coordination of a variety of knowledge structures and linguistic abilities. [10]

For children who use AAC, narrative competence carries particular significance and particular risk. Children who rely on AAC systems may face particular challenges in developing narrative competence. These difficulties are not solely attributable to underlying language impairments, but also to the restricted opportunities for language learning that AAC users often encounter, as well as the structural and expressive limitations inherent to many AAC systems themselves. Research has documented that the narrative discourse produced by children who use AAC tends to lack organization and coherence, and is frequently limited in both the range of vocabulary and the complexity of grammatical structures employed [11]. Furthermore, Children who use AAC frequently depend on their communication partners to co-construct narratives, often taking a supporting rather than leading role in the storytelling process. This dynamic, while sometimes necessary, can result in limited opportunities for the child to contribute narrative elements autonomously, potentially restricting the development of independent narrative skills over time [11].

This gap between communicative potential and actual narrative opportunity has motivated a growing body of research on AAC-based narrative intervention. However, a substantial body of research indicates that the overwhelming majority of AAC interventions documented in the literature have prioritized requesting as the primary communicative function, with some reviews estimating this proportion at between 60 and 90 percent. While the development of **requesting skills** constitutes a necessary foundation, it represents only one dimension of communicative competence. Functions such as labeling, commenting, and sharing information about experiences and events are equally important, as they underpin social engagement, the development of interpersonal closeness, and the ability to convey and exchange meaningful content with others [12].

Storytelling activities have shown particular promise as a vehicle for AAC intervention. Research has demonstrated that integrating AAC into dramatic storytelling contexts can support the acquisition of functional communication vocabulary and increase the frequency of verbal communication attempts. The group dimension of these activities adds a further layer of value, as shared storytelling creates naturally occurring opportunities for peer imitation alongside more structured and intentional AAC instruction [13]. This finding is relevant beyond the clinical setting: it suggests that narrative is not merely a target for intervention, but a context that motivates and amplifies communication development more broadly. [13]

The book *Using Storytelling to Support Children and Adults with Special Needs* [14] further establishes storytelling as a practical therapeutic and educational tool, one that can be adapted to diverse cognitive profiles, communication modalities, and support contexts. This perspective aligns directly with the design premise of AACreative: that the act of constructing a story, mediated by physical symbol cards and generative AI, is not a supplementary feature but a core therapeutic and developmental activity in its own right.

2.5 Tangible User Interfaces

"Coupling Bits and Atoms"*[15]*

Building on this theoretical foundation, Hiroshi Ishii and Brygg Ullmer at the MIT Media Lab proposed the concept of Tangible User Interfaces (TUI) as a concrete design paradigm. Their vision: "Tangible Bits"[15], sought to bridge the gap between cyberspace and the physical environment by making digital information tangible, allowing users to grasp and manipulate bits by coupling them with everyday physical objects. The goal was to establish a new type of Human Computer Interfaces (HCI) that augments the real physical world by coupling digital information to everyday physical objects and environments.

This represented a fundamental departure from the dominant GUI paradigm. Where graphical interfaces rely almost entirely on transient visual representations on screens, “tangible interfaces give physical form to digital information, employing physical artifacts both as representations and controls for computational media”[16]. TUIs couple physical representations, spatially manipulable physical objects, with digital representations such as graphics and audio, yielding interactive systems that are computationally mediated but generally not identifiable as computers per se[16].

Ullmer and Ishii formalized this distinction through the MCRpd interaction model (Model, Control, Representation -physical and digital-) which they proposed as an alternative to the GUI's conventional MVC framework (Model-View-Controller). In the MVC model, input and output are clearly separated: the controller, operated through peripherals such as a keyboard or mouse, is distinct from the view, which displays information on a screen (Figure 3a). The MCRpd model, by contrast, highlights a defining characteristic of TUIs: the physical artifact simultaneously serves as both the control mechanism and the representational surface of the interface (Figure 3b). This integration is not merely conceptual, it is materially present in the object itself, which carries information and enables interaction through the same physical form, collapsing the separation between input and output that graphical interfaces maintain[16].

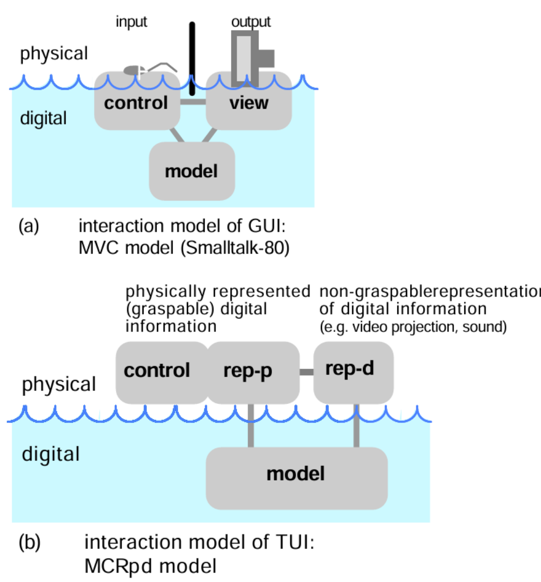


Figure 3a,b. GUI and TUI interaction models. [16]

A critical property of TUI systems is the persistence of physical state. Unlike graphical elements that can be created or destroyed with a button press, physical tokens remain in the world and can be read both by people and computers through their configuration. This fourth key characteristic of tangible interfaces means that their state is coupled with what they represent in the digital state [16]. In the context of AACreative, this property is directly instantiated by the NFC cards: each physical card persistently embodies a digital association, readable by both the child choosing it and the sensor detecting it.

It is also worth noting that Ishii and Ullmer themselves identified RF-ID tag technologies as among the most promising technical mechanisms for realizing physical-digital couplings in TUI systems. Over the last decade, the use of wired and wireless ID tag technologies has grown increasingly popular, with RF-ID tags showing special promise as a sensing approach for tangible interfaces.[16]

On the other hand another key comparison is that GUI interfaces tend to lean heavily on cognitive skills, overshadowing other skills while Tangible Interaction can ‘open up skills through skills’, in this case cognitive, linguistic, social and emotional skills through perceptual-motor skills (figure 2).[17]

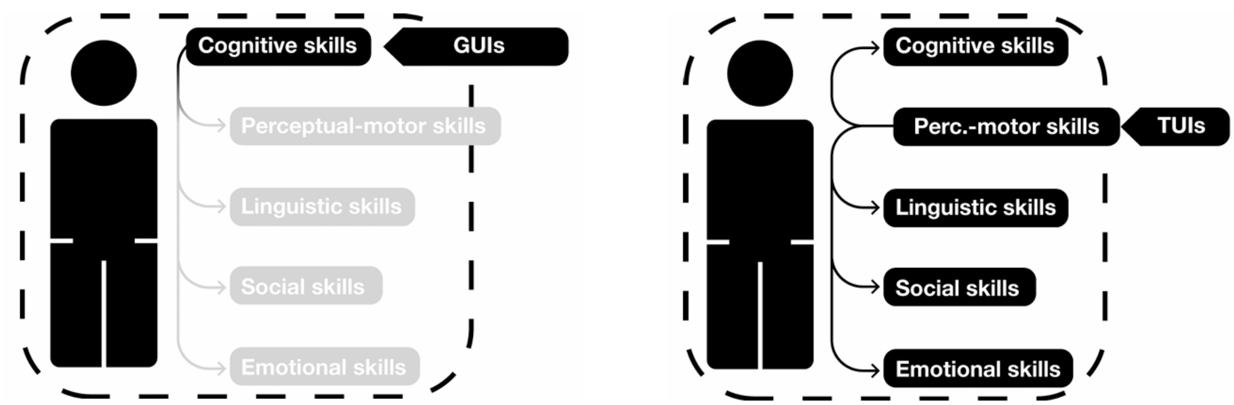


Figure 4. a. (left) GUI-based interfaces lean heavily on the child's cognitive skills. b.(right)TUIs open up skills through perceptual-motor skills. [17]

Digital Manipulatives: TUI in the Service of Learning

Within the broader TUI paradigm, a specific strand of research has focused on the pedagogical potential of physical-digital objects in children's learning contexts. Resnick et al. (1998) introduced the concept of *Digital Manipulatives*, computationally augmented physical objects designed to support learning through hands-on manipulation, extending Piaget's constructivist tradition of physical manipulatives (which explains how through the manipulation of physical objects like blocks, rods or spheres, abstract concepts like number, quantity, space or causality can be understood) into the digital age[18]. Several systems in the constructive and relational TUI categories have been oriented towards use in educational contexts; researchers have highlighted two particularly relevant affordances of physical artifacts in this context: their capacity to support simultaneous interaction among multiple users, facilitating collaborative engagement around a shared object, and their ability to make abstract concepts tangible through concrete, manipulable representations, a quality that has been recognized as especially valuable in educational and learning environments[16].

This pedagogical framing is particularly relevant for AACreative. The NFC cards function as digital manipulatives in the precise sense Resnick intended: they are not merely input devices, but physical objects whose manipulation carries meaning. When a child reads a card, the act of choosing it is part of the communicative and narrative act — the physical gesture and the digital outcome are inseparable.

Relevance to the AACreative Device

The framework presented in this section, TUI as design paradigm, and digital manipulatives as pedagogical application, provides the theoretical basis for the tangible device developed in this thesis. Tangible interaction was introduced as a mixed discipline integrating the physical with the digital; yet from an embodied perspective, the most successful designs are those that carefully integrate physical and digital worlds while leaving the physical world alone to the extent possible, admitting the improvisations of practice that the physical world offers.

The AACreative device operationalizes this principle. Rather than replacing the familiar AAC symbol card with a purely digital interface, it preserves the physical card as a meaningful, graspable artifact while extending its communicative potential through digital augmentation. The result is an interaction modality that is at once theoretically grounded, pedagogically appropriate, and practically accessible for the children who use it.

3. LITERATURE REVIEW 

3.1. CASE STUDY ANALYSIS

Personalized storybooks using generative AI: Autihero

AutiHero is a generative AI system designed to support parents of autistic children in creating personalized **social narratives** for behavioral guidance. By incorporating each child's specific interests, familiar people, and everyday contexts, the system generates custom stories with AI-produced illustrations, making social learning more relatable and contextually grounded. The platform features a branching story structure that allows children to explore the consequences of different behavioral choices, including undesirable paths that lead to repair opportunities and positive resolutions. Personalization operates across three dimensions: **persons, places, and interests**, which parents input directly, while the target behavior to be addressed is also specified by the caregiver. A two-week deployment study with 16 parent-child dyads demonstrated that the system effectively reduced the cognitive burden on caregivers while fostering active participation and observable behavioral changes in children. [19]

The story generation pipeline incorporates explicit quality controls, including reading level assessment using the Flesch-Kincaid Grade Level scale (a readability formula that assigns a score equivalent to a US school grade level based on sentence length and word complexity [20]) and vocabulary difficulty evaluation based on the CEFR framework (the Common European Framework of Reference for Languages, which classifies language proficiency across six levels from A1 beginner to C2 mastery [21]), ensuring that output is appropriate for the child's developmental stage. Prompt design also plays a central role in narrative coherence: the system uses structured instructions that include, among others, the requirement to "use sentence structures and vocabulary that are appropriate for elementary school students" [19], maintaining consistent character appearance across illustrations, and restricting emotion vocabulary to a predefined word list appropriate for young children.[19]. Figure 5 shows the personalization options in the Autihero User Interface.

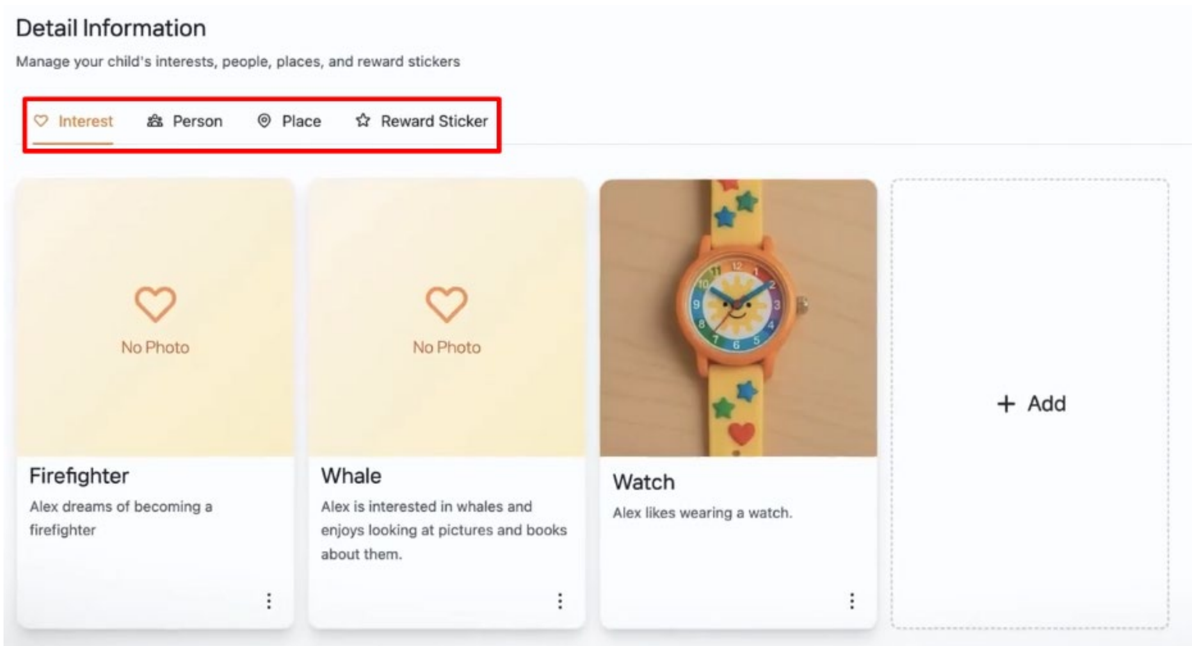


Figure 5. User Interface of Autihero. In this section the caregiver gives details about the child. [19]

"NarrAid: Supporting Storytelling of People with Aphasia via Generative Visual Scene Displays"[22]

NarrAid is an innovative assistive system designed to help **people with aphasia** (Inability to comprehend or formulate language) share personal stories through **generative visual scene displays (VSD)**. Traditional communication tools for those with language impairments often focus on basic needs and require heavy manual preparation by caregivers, but NarrAid utilizes **generative artificial intelligence** to transform fragmented words or image selections into **high-context scene images**. By offering **structured narrative guidance**, focusing on elements like time, place, and emotion, the system allows users to co-create vivid visual anchors that facilitate **shared understanding** and reduce social isolation. Results from a preliminary user study indicate that this AI-assisted approach significantly increases the **confidence and agency** of people with aphasia while lowering the practical burdens of content creation for their support networks. Ultimately, this tool illustrates how **multimodal interaction** and AI can move beyond simple speech generation to support the more complex, human need for **semantically rich self-expression**.[22]

NarrAid has two input methods, by recording voice or by selecting ARASAAC symbols [3] organized in six categories: **Time, Location, People, Actions, Objects, Emotions (Mood).** The system uses a **structured guidance mode** (also a free style one) that proactively asks the user questions related to these elements (e.g., "Where did this story take place?"), and the user responds by selecting the corresponding ARASAAC cards [3], [22].

Once the user selects a series of pictograms, the system's **Generative AI** integrates these fragmented inputs into a coherent narrative representation. This process works as follows:

- **Input Processing:** The system takes the selected ARASAAC labels (e.g., "garden," "father," "planting").
- **Prompt Generation:** These labels are combined with any voice input into a detailed prompt.
- **Output:** The AI generates a **high-context scene image** and corresponding story text that reflects the chosen pictograms, creating a Visual Scene Display that serves as a shared point of reference during storytelling.[22]

The interface includes an input bar where selected cards appear, and it allows caregivers to **upload personalized images** that can be used alongside the standard ARASAAC library[3]. Qualitative insights from the study indicate that users often prefer selecting these image cards over speech input, as it aligns better with their ability to provide one- or two-word answers to structured questions. Figure 6 explains the NarrAid User Interface.[22]

Figure 6. “Interface of NarrAid. (A) AI-generated guidance. Here ($a1$) is Mode-switching bar between free-chat mode and guidance mode. (B) The Input module includes: ($b1$) input bar with delete button, ($b2$) image card selection (pictographic symbols are sourced from ARASAAC), ($b3$) core buttons including Yes, No, Skip, Not clear, and Generate, and ($b4$) custom image card component allowing caregivers to upload personalized cards which then appear in ($b2$). (C) The window to display high-context images generated by AI. (D) My Stories archive to store generated images, texts, and cards in a card-style layout” [22]

“Moovy: a Tangible User Interface to Train Linguistic Skills in Children”[23]

Moovy is a Tangible User Interface (TUI) designed as a tabletop game to help children with **Developmental Language Disorder (DLD)** master complex grammatical structures, specifically **clitic pronouns** and **passive sentences**. The system translates abstract linguistic concepts into concrete physical actions.[23]

The Moovy system consists of an **interactive board** (equipped with RFID readers and LEDs), a **table-top panel** representing game boxes, **Tangible Argument Objects** (RFID-tagged toys or cards), and a **web app** for caregivers to manage sessions. The gameplay follows a specific multisensory loop:[23]

- **Audio Stimuli:** The board plays a pair of spoken sentences (e.g., "First the child eats the cake. Then the cake is eaten by the grandfather").

- **Argument Identification:** The child must identify the "arguments" (the participants in the action, like the child or the cake) from the physical objects on the table.
- **Physical Manipulation:** The child places the objects in specific cells on the board's 4x2 grid to represent their roles in the sentence.
- **Feedback:** The system provides immediate reinforcement; **green LEDs** indicate correct placement, while **red LEDs** indicate errors. Successful sessions conclude with a "rainbow light" effect and cheerful music.[23]

The core innovation of Moovy is the use of **physical interaction** to facilitate "argument movement": the mental displacement of a word from its usual position. So, the physical interaction allows children to “make real” invisible grammatical operations. By physically moving a toy from one cell to another, the child enacts the syntactic transformation occurring in the sentence. The game uses the **"Ghost Trace" Effect,** to reinforce the concept of movement, by illuminating the cell an object just vacated with a white light, visually marking the transition in the sentence structure. The game offers a **Neural Overlap between Language and Tool-Use:** The research is based on neuroscientific evidence showing that **tool manipulation** and complex **syntactic processing** share the same neural substrates in the **basal ganglia**. Meaning that engaging in intentional, goal-directed physical manipulation recruits these shared circuits, potentially fostering a **learning transfer** from motor actions to linguistic skills. Besides, **f**ollowing the principles of **Embodied Cognition**[24], Moovy assumes that language comprehension is enhanced when linguistic symbols are mapped onto **action-based simulations**. The physical engagement helps synchronize cognitive processes with bodily experiences, making the learning more robust and enjoyable.[23]

Experimental results showed that children using Moovy **doubled their production** of correct target sentences (specifically clitic pronouns) after a six-week program, whereas a control group receiving traditional therapy showed no significant improvement.[23]

Linguabytes: Tangibles for toddlers learning language

Linguabytes is a modular, tangible play-and-learning system designed in 2011 to stimulate the language and communication skills of **toddlers (developmental age 1–4 years)**

with multiple disabilities. Using a **Research-through-Design (RtD)**[25] approach, the authors created a system that scaffolds a child's learning within their Zone of Proximal Development (defined as the range of tasks that a developing child is unable to perform independently but can accomplish with the **assistance and shared experience of more capable peers or adults**) [17] through interactive storytelling and targeted exercises.

System Synthesis: The LinguaBytes Prototype

The system is composed of five modular parts designed to be flexible and easily adaptable for therapy sessions:

- **Central Console:** A 17-inch screen and speakers providing multimedia feedback.
- **Base Module:** The core interface for the child, featuring RFID readers and a **scenery-slot** that acts as a "tangible programming" tool to select themes (e.g., animals, vehicles).
- **Interface Modules:** Swappable modules tailored to specific linguistic skills: the **BookBooster** for interactive reading, the **Phonotron** for phonological awareness, the **PuzzlePointer** for semantics, and the **Combination Module** for building 1- to 3-word sentences.
- **Carer's Controller:** A dedicated physical remote allowing parents or therapists to pause, repeat, or toggle between explorative and assignment-based modes without using complex menus.
- **Tangible Input Materials:** A collection of RFID-tagged storybooks, figures, and symbols, including **programmable RFID labels** that can turn a child's own toys into system inputs.

Relevance of Tangible User Interfaces (TUIs)

The authors argue that traditional Graphical User Interfaces (GUIs) rely too heavily on cognitive skills, which can overshadow other developmental areas in young children. They highlight several key reasons why **Tangible Interaction** is superior for this population:

- **Opening "Skills through Skills":** TUIs allow children to utilize their **perceptual-motor skills** to unlock development in cognitive, linguistic, social, and emotional domains. For children with **Cerebral Palsy**, enhancing motor interaction directly leads to a richer learning environment and higher self-esteem.
- **Lowering Cognitive Load:** By capitalizing on **real-world affordances** (objects that children naturally understand how to grab and move), the system reduces the cognitive effort required to interact with the technology, allowing the child to focus on language acquisition.
- **Slowing Down Interaction:** Unlike fast-paced digital media, physical interaction **naturally slows the pace of the activity**. This gives children more time to process information and gives carers better control over the timing of the session.
- **Non-Disruptive Adaptability:** In speech therapy, time is limited (often 30-minute sessions), and therapists are rarely "the programming kind". TUIs allow for **immediate, physical changes** to settings (like swapping an interface module or sliding in a new scenery) without the need to navigate disruptive digital menus.
- **Fostering Initiative and Expression:** Testing showed that the tangible nature of the system encouraged children to take more **communicative initiatives**, resulting in longer attention spans and more frequent verbal or gestural expressions. It also makes the system highly suitable for **collaborative use** between the child, the caregiver, and potentially other peers

3.2. BENCHMARKING

3.2.1. AAC APPS

The following section analyzes several applications used by AAC users for communicative purposes; they were selected because they were mentioned by the experts or because they appeared referenced in some of the sources consulted. The other applications reviewed were consulted because they were mentioned by experts as being free to access or offering a free trial, and because they appear as associated recommendations on the Play Store.

AVAZ

Avaz is a picture- and text-based AAC application designed for children and adults with complex communication needs, including autism, cerebral palsy, Down syndrome, aphasia, and apraxia. It operates on iOS and Android tablets and smartphones.[26]

The app is structured around a vocabulary grid where users tap pictogram buttons to build sentences, which are then read aloud via text-to-speech. It supports multiple vocabulary levels, allowing progression from single words to full sentences, and includes a keyboard mode for more advanced users. [26]

Avaz's navigation logic is based on a grid that places core words on the main screen, organized from left to right by categories such as: quick-access words, pronouns, verbs, connectors, prepositions, and descriptors that modify a word in a sentence. It also features a row of folders at the bottom, categorized as: **Quick Access, People, Questions, Actions, Feelings/Sensations, Time, Short Words, Places, Descriptors, and Special Interests**. Additionally, it includes a column of folders covering topics such as: **Things, Body Parts, Food, Transportation, and Topics**. These folders lead to further folders or to their associated words, maintaining a color-coding system for grammatical categories. Beyond the grid containing the words, the app includes a top bar for building sentences and converting them into speech using a specific reading tone (angry, excited, question, sad, sarcastic, scared, whisper).[27]. Figure 7 shows the User Interface of the App in the home page.

Among its customization options, the application offers resources such as: recording specific audio for each word, uploading images from the device or from the internet, or taking photographs directly within the app, adding GIFs or YouTube video segments, and configuring favorite phrases that are organized in a quick-access folder.[27]

Symbol system: Avaz uses the SymbolStix pictogram library, offering over 40,000 symbols. Caregivers can also add custom images. [26]

Color system: Avaz uses the Fitzgerald Key color-coding system (a system that groups pictograms into six colors depending on their grammatical role, besides the system proposes organizing the pictograms from left to right according to their role [28]), which makes it consistent with other special education materials in the child's classroom. [26]

Figure 7. User Interface of AVAZ, App for communicating using pictograms.[26]

Proloquo2Go

Proloquo2Go® is a symbol-based AAC app designed for people who cannot speak or have difficulty making themselves understood [29]. It utilizes the Crescendo™ vocabulary (developed specifically for that solution) that supports multiple grid sizes and three levels of language skills. Proloquo2Go also provides a keyboard view with **word prediction** for users with literacy skills. From the accessibility perspective Proloquo2Go facilitates methods like scanning and Select on Release. [29]

It is designed for children and adults who cannot speak or need support to be understood, including individuals with autism, cerebral palsy, Down syndrome, and Angelman syndrome. The app is available exclusively on iOS.[30]

Its core navigation is built around the Crescendo™ vocabulary system, a research-based structure that places core words in consistent locations to support motor planning and language growth, allowing users to progress from single words to full sentences within the same application. Crescendo offers three pre-programmed vocabulary levels — Basic Communication, Intermediate Core, and Advanced Core — and users can move between levels while retaining their customizations. The grid can display between 9 and 144 buttons per page, accommodating a wide range of motor and cognitive profiles. [30]

Navigation follows a folder-based structure: tapping a symbol either adds it to the message bar or opens a subcategory folder (see Figure 8). Words added to the message window can then be spoken aloud, and the app supports secondary actions such as verb conjugations accessible through long-press or double-tap gestures. [29]

Symbol system: Proloquo2Go uses the SymbolStix symbol library. SymbolStix uses stylized stick figures to depict people in its images. The app includes more than 27,000 symbols, and caregivers can also add custom photos taken within the app or selected from the photo library. [30]

Color system: Proloquo2Go's Crescendo vocabularies use the Modified **Fitzgerald Key**, a color-coded display system designed to increase visual access and reinforce linguistic relationships between grammatical categories[30]

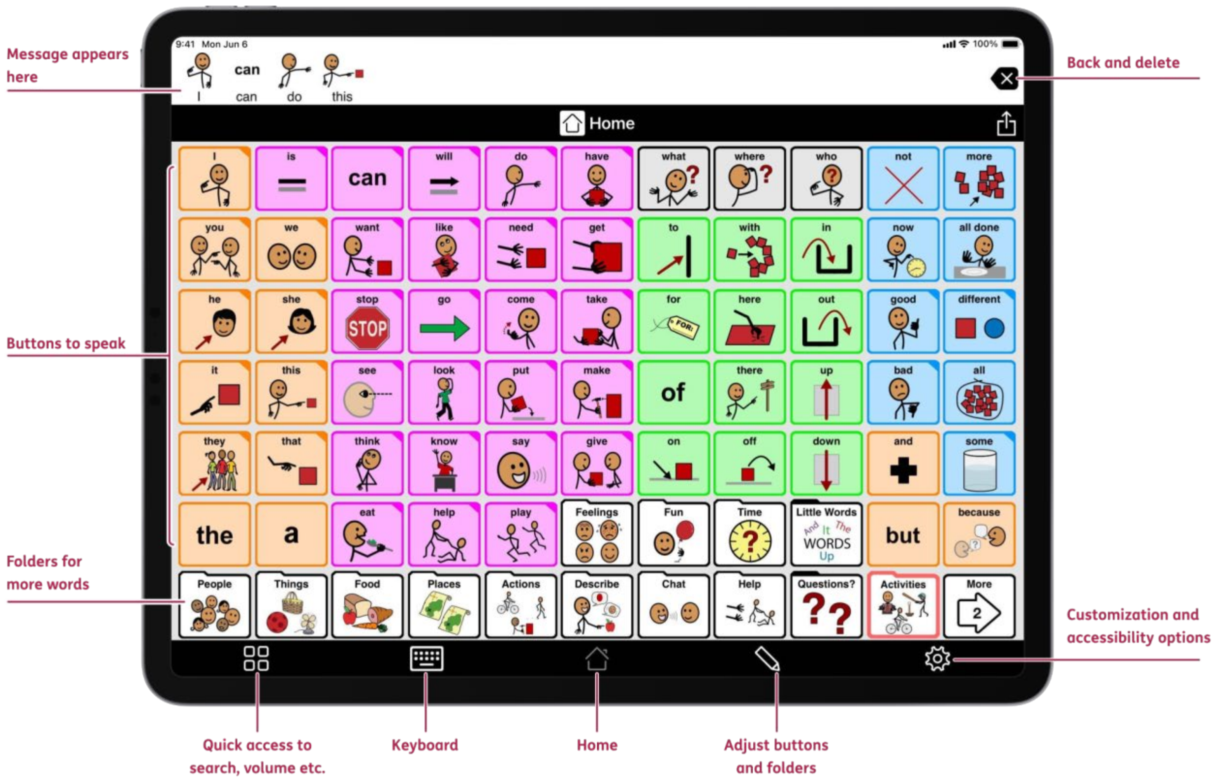


Figure 8. Proloquo2Go User Interface and its functionalities [30].

Other less popular and more recent apps consulted were **CommBoards AAC, Otsimo AAC, Larunda**, and **dictaPicto** (Figure 9)**.** The first two follow a logic of grouping words into thematic folders, as well as using a top bar to build sentences that can be read aloud by the application. Larunda, on the other hand, prioritizes the creation of boards, groupings of words by topic, and includes predefined boards with core vocabulary of 20 or 50 pictograms, where words can be accessed directly; these words are distinguished by color-coded categories. Its base pictograms come from the ARASAAC library, but it also allows personalization by uploading images from the gallery or taking photographs. Within the boards, it distinguishes between fixed and main pictograms: the former are accessible through a horizontal bar, while the latter are accessed via vertical scroll, and both can be modified. Finally, dictaPicto takes a different approach: it stores phrases in thematic folders, phrases that are composed using keyboard or voice input. This indicates that in this application the caregiver plays a fundamental role in initiating the configuration of these phrases, which can also be personalized by replacing symbols with photographs taken or uploaded from the device.

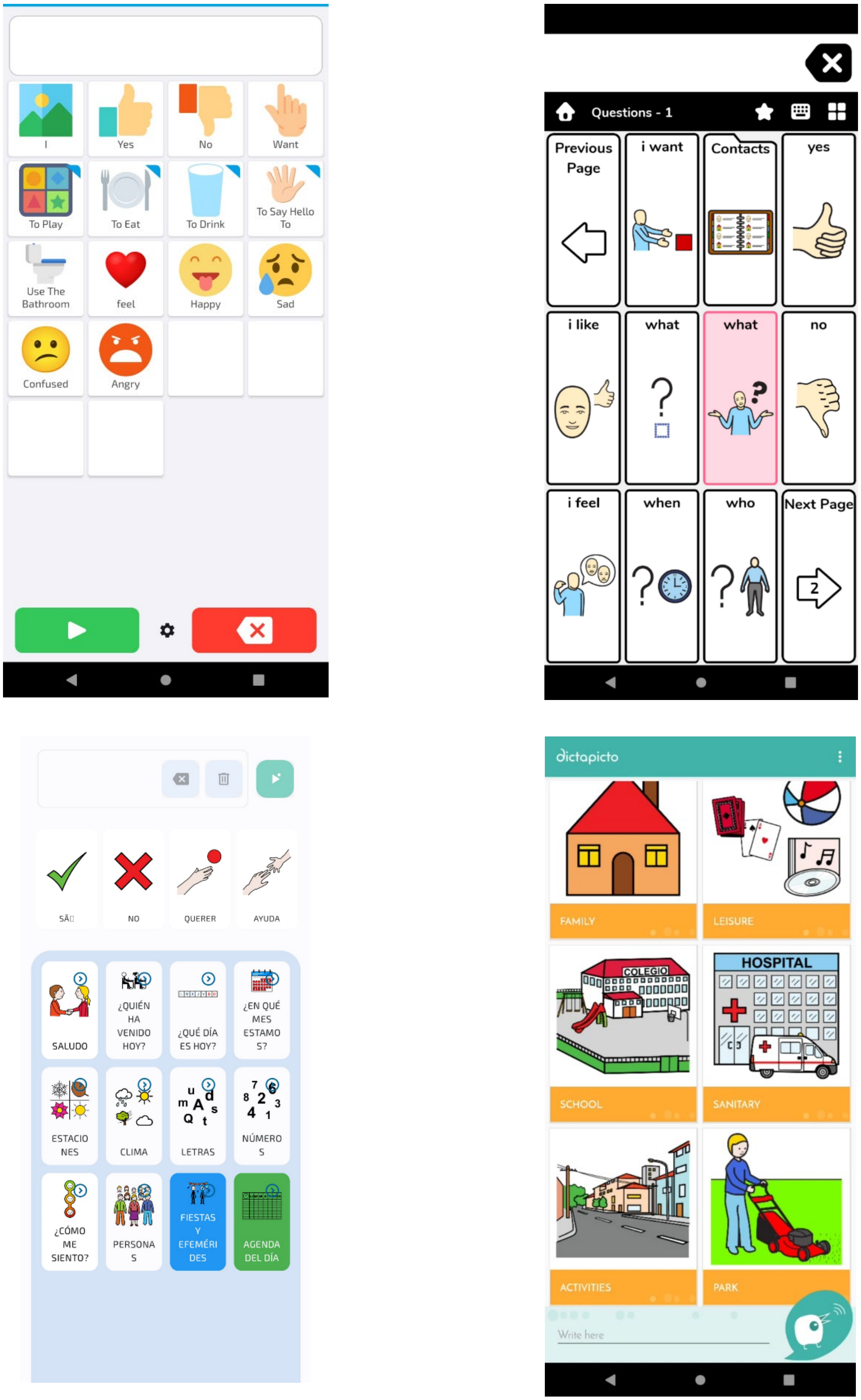


Figure 9. Other Apps explored: Left to right, top to bottom: Commboards AAC, Otsimo AAC, Larunda and DictaPicto.

4. PROJECT DEFINITION

4.1. AACREATIVE: THE STARTING POINT

AACreative is a tablet application developed as a final project for the Advanced User Interfaces course at Politecnico di Milano during the academic year 2025/2026, by a multidisciplinary team of four students: Daniella Laloo, Danyil Voytekho, Mattia Rainieri, and Elisabetta Vanoli. The application was conceived as a cooperative storytelling tool for children with cognitive and speech impairments, combining symbol-based AAC interaction with generative artificial intelligence to produce **personalized multimodal storybooks**. The app is built using React Native with the Expo framework, making it cross-platform and deployable on tablet devices running both iOS and Android. The backend is hosted on Microsoft Azure and integrates with OpenAI's API for both text and image generation.

The application is structured around three main sections, accessible through a bottom navigation bar: Caregiver, Create, and Library.

Caregiver section

This section is designed for the caregiver that supports the AAC user. Here, they assemble sets of up to twelve pictogram cards that will constitute the selection framework available to the AAC user to create the story. The idea of having these sets is to ease the selection process for the child decreasing the mental overload of choosing among thousands of pictograms; instead the caregiver narrows the approach grouping the cards by an interesting topic or something that he/she wants to address (for instance a therapist or a parent can have the interest of speaking about a specific topic and doing that through a story can facilitate the process). These cards can be browsed through category filters or located via a search bar (Figure 10), drawing from the ARASAAC symbol library [31]. Once selected, the cards are grouped into a named set, which is then added to a collapsible overlay menu on the left side of the interface (Figure 12). An additional option, accessible through a "Created by me" filter (see Figure 10), allows caregivers to upload or capture images to create custom cards (Figure 11), which can then be incorporated into any set alongside standard ARASAAC pictograms. This feature is fundamental so the child can recognize familiar elements instead of abstract symbols given him/her a bigger sense of identification.

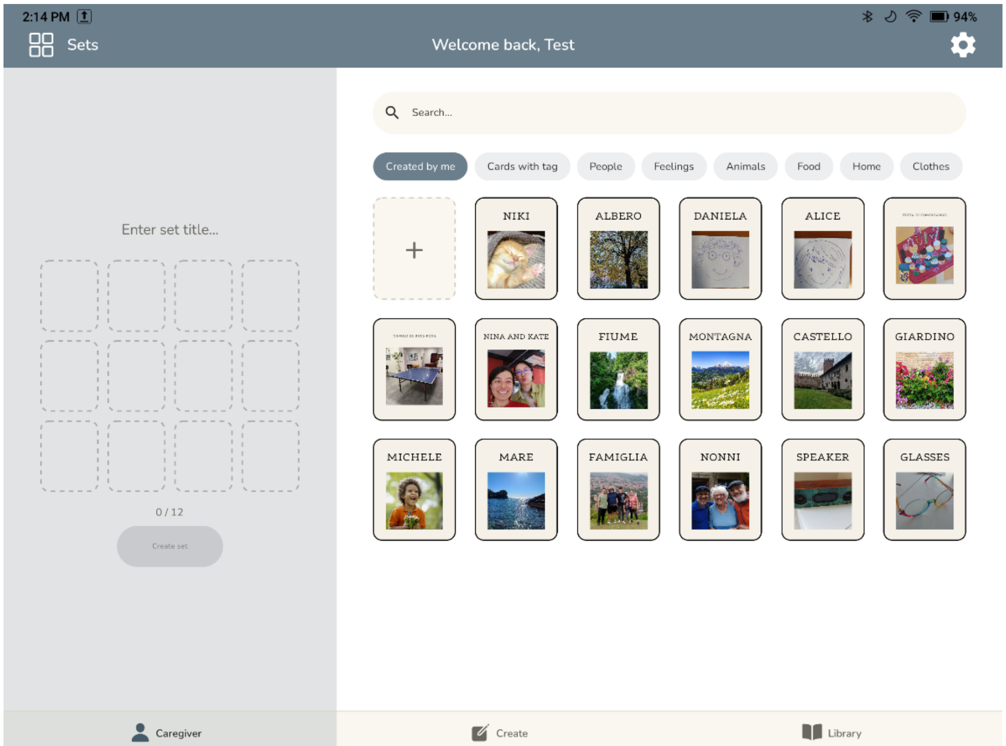


Figure 10. AACreative app: Caregiver section

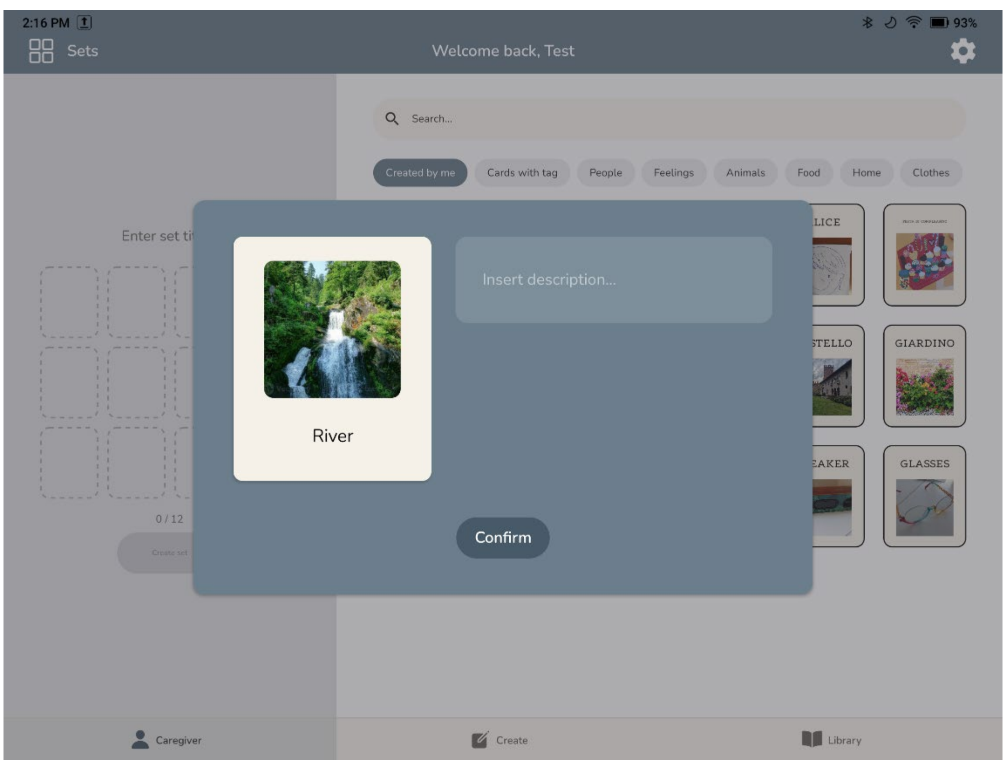


Figure 11. AACreative App: Creation of a personalized card

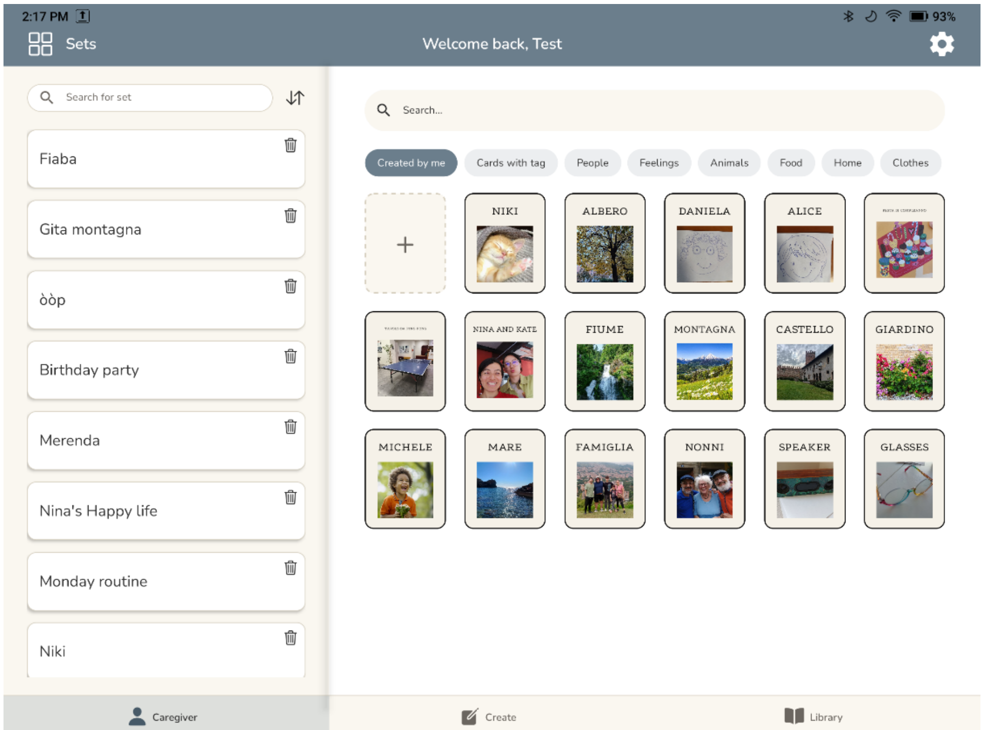


Figure 12. Overlay with the sets created by the caregiver.

Create section

This is the primary interaction space for the AAC user. Once the caregiver has selected the active set, the AAC user can browse the available cards and select those they wish to include in their story (Figure 13). Tapping a card triggers audio playback of the word it represents, providing immediate multimodal feedback. When the child presses the Create button, the selected cards are sent to the server, where the LLM generates a multimodal story based on the symbolic input. During generation, a waiting screen displays the selected cards to maintain the child's engagement (Figure 14). The resulting storybook consists of three to four pages, each presenting an AI-generated illustration alongside a short text passage and a symbolic representation using AAC pictograms (Figure 15). The story can be listened to through text-to-speech narration and navigated page by page in both directions.

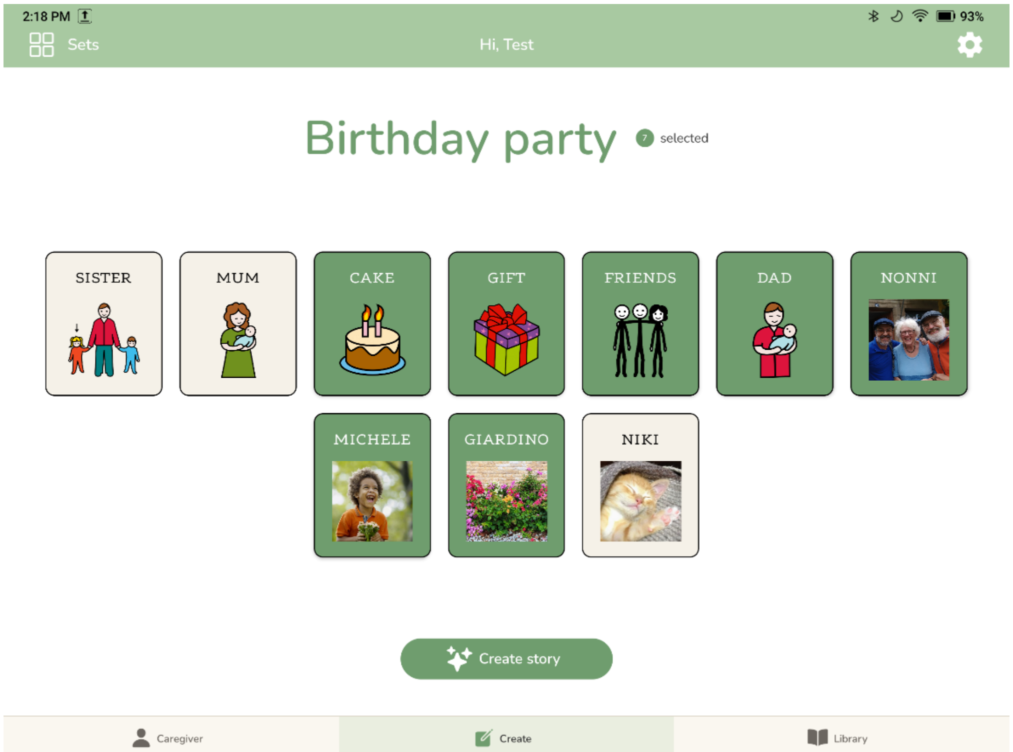


Figure 13. AACreative App: Create section, designed for the AAC user

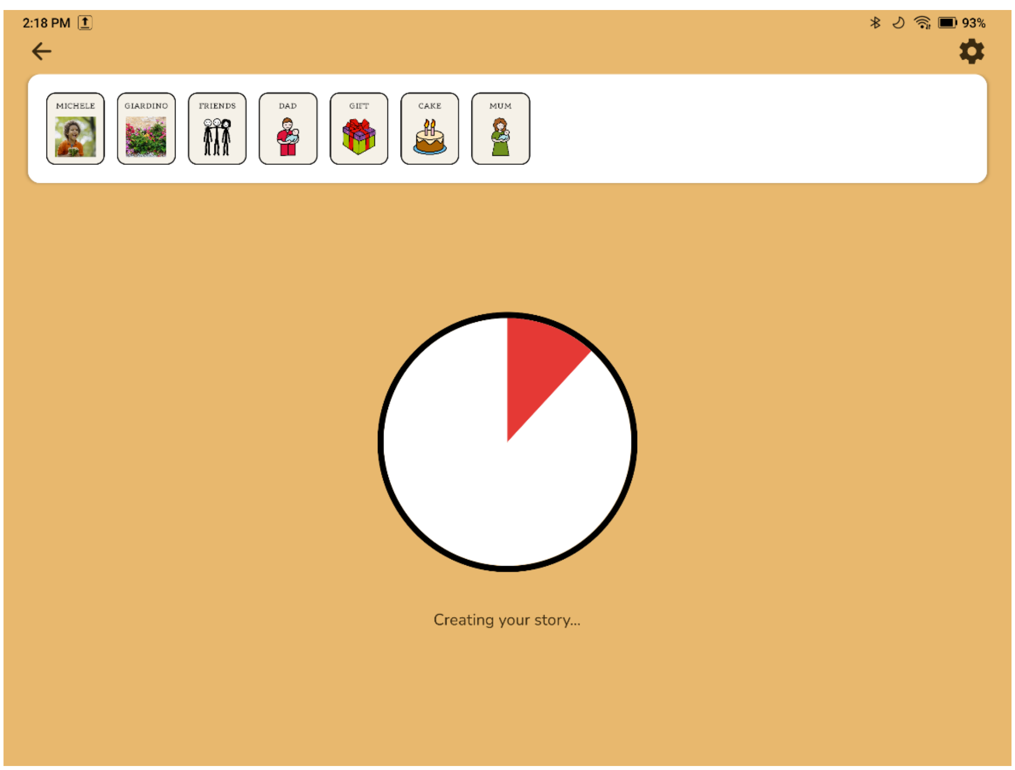


Figure 14. AACreative: Screen for waiting the generation of the story.

Figure 15. AACreative: Going through the multimodal story

Library section

All generated storybooks are stored in this section, organized chronologically from most to least recent, allowing caregivers and children to revisit previously created stories (Figure 16).

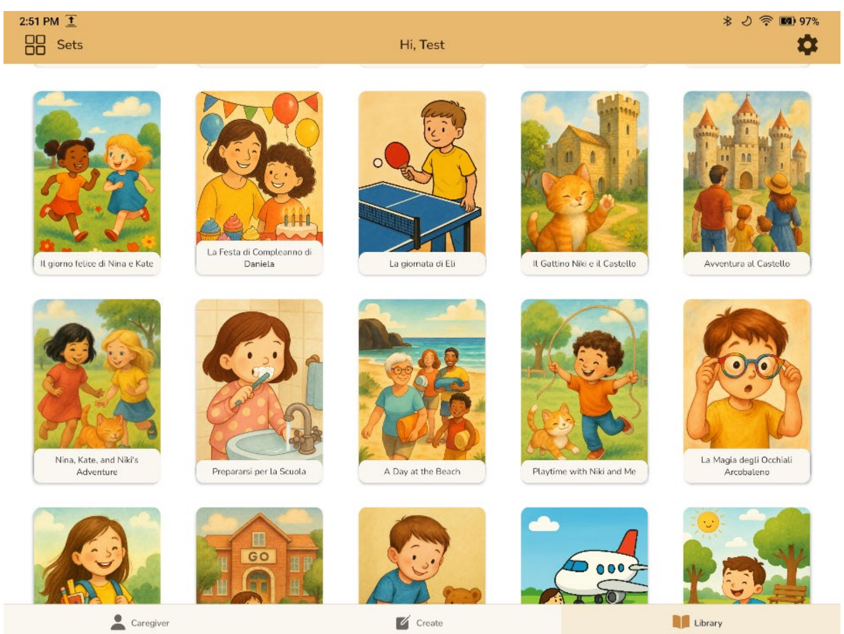


Figure 16. AACreative: The Library section

In terms of configuration, the application currently allows audio playback of card labels to be toggled on or off. It also includes the login and logout screens needed to keep separate user profiles, so each one has its own personalized card sets and story library adapted to the specific needs of the individual AAC user.

4.2. UNDERSTANDING THE CONTEXT OF USE

4.2.1. Exploratory Digital Content Review

Prior to conducting formal interviews, an informal review of digital content was carried out on YouTube, Instagram and Pinterest, with the aim of gaining an initial understanding of how AAC systems are used in everyday, non-clinical contexts. This exploratory phase helped orient the subsequent interview process by identifying recurring practices, common tools, and the vocabulary used by the AAC community. The content encountered was produced primarily by parents, speech-language pathologists, and inclusion teachers, who use these platforms to document their experiences and share strategies with others navigating similar contexts; the most common encountered material showed autistic minimally verbal children.

The material reviewed revealed that AAC is most commonly used for two communicative functions: **requesting and routine planning**. Videos frequently depicted caregivers preparing and organizing pictogram-based materials, folders, portable accessories, door-hanging agendas, demonstrating the considerable time and effort involved in making AAC systems consistently available across different environments. Parents in particular, showed creative solutions for portability: printed pictograms attached to small wearable accessories, lanyards, or custom-built holders designed to always keep the system accessible. Several caregivers also demonstrated **personalization strategies**, incorporating photographs of familiar places, people, and objects alongside standard pictograms, confirming that generic symbol sets are routinely adapted to reflect the child's individual context.

A recurring feature of the content was the emphasis on **repetition** as a learning mechanism. Caregivers consistently modeled the same communicative act multiple times to reinforce its association with a desired outcome. In videos showing PECS (Picture Exchange Communication System [32]) implementation[1], for example, the exchange ritual was repeated across multiple trials: the child selects an image representing a desired item, hands it to the caregiver, and receives the item as reinforcement[1]. In this system, children were asked not merely to point to a symbol but to physically detach it from a board and place it in a designated area, introducing a manipulative, tactile dimension to the interaction. In this setup, Velcro was a common used material, to facilitate rearrangement of the pictograms by the AAC user or by the caregiver.

Observations of device use also revealed relevant usability patterns. In several videos, AAC users did not maintain consistent visual attention on the tablet screen, and in some instances pointed indiscriminately at symbols, causing the application to repeat the same word multiple times in succession. This suggests that sustained attention and precise motor targeting on touchscreen interfaces present genuine challenges for some users, a finding that aligns with clinical accounts gathered in the interview phase. Additionally, caregivers noted that many of the most feature-rich AAC applications, such as Proloquo2Go[2] and LAMP Words for Life[3], are available exclusively for Apple devices and require paid subscriptions, creating accessibility barriers for families with limited resources. Physical device protection also emerged as a practical concern: AAC communicators were frequently shown with reinforced casings and large grip handles, reflecting the need for durability and ease of handling in everyday use.

[1] Youtube video: Communication Throughout The Years by ROCK'N STIMS:
https://www.youtube.com/watch?v=I5MsG0dTxqw
[2] Youtube video: AAC App comparison by Lillys Voice Inc:
https://www.youtube.com/watch?v=FuVLmA2b9Z0
[3] Youtube video: LAMP words for life tour by Casi Heartstrong, CCC-SLP, AT:
https://www.youtube.com/watch?v=bTm8QN8hlhg&list=PLLLCDy98fG2cnAYv6spnQNJaexNNrMBqe

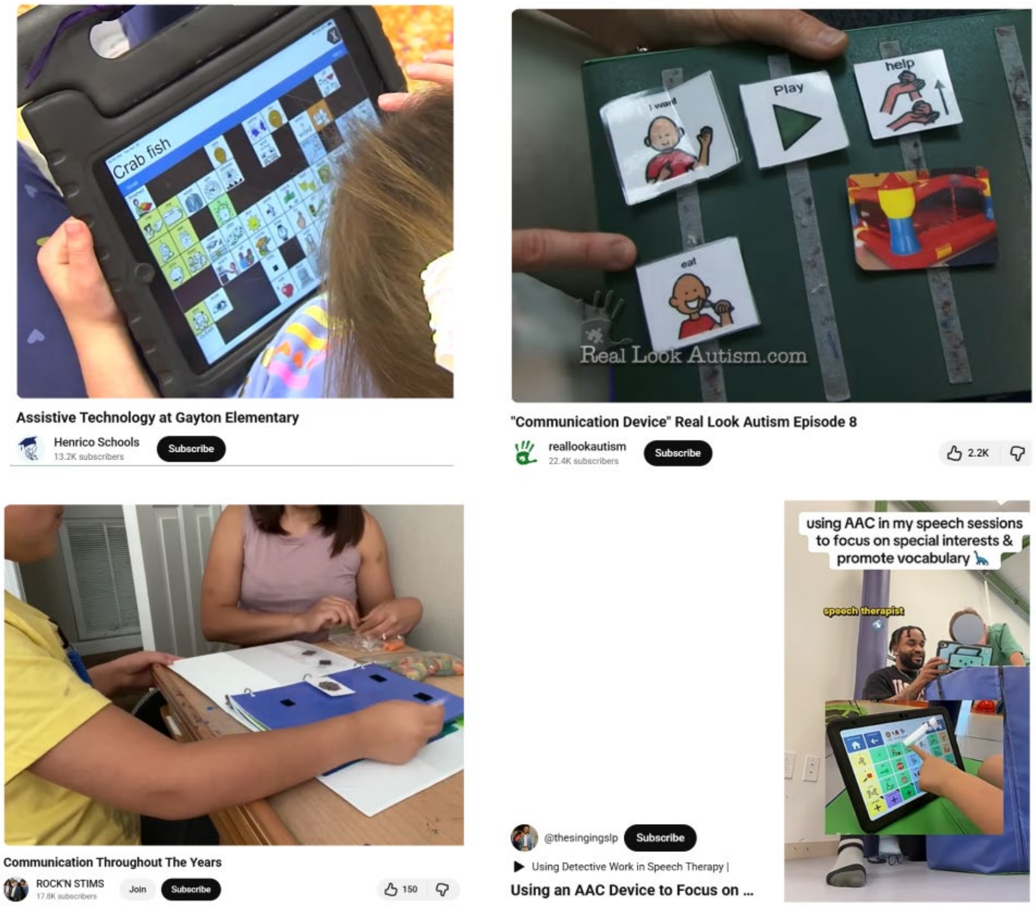


Figure 17. Example of AAC used in different settings. Source: Youtube

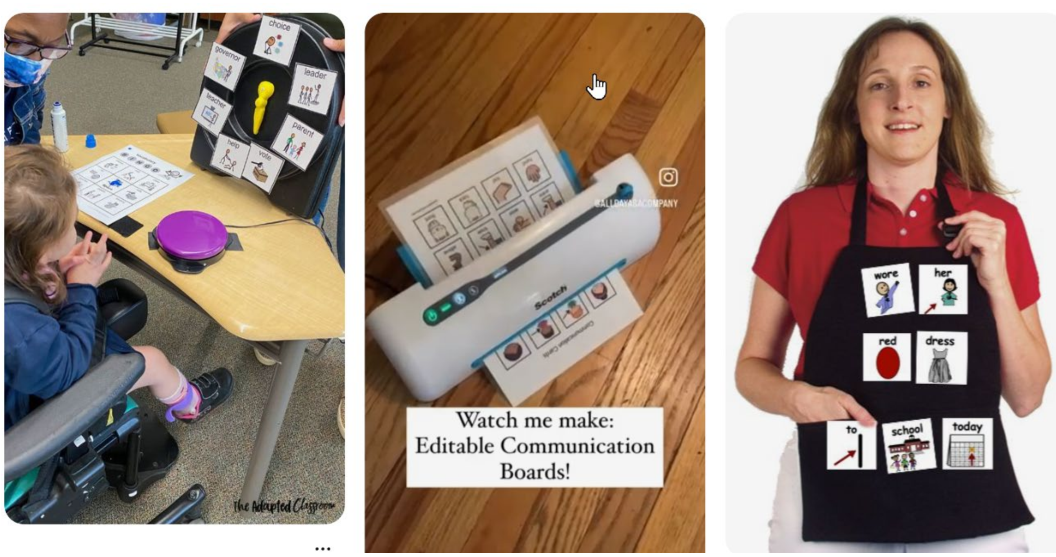


Figure 18. Examples of use and preparation of AAC pictograms using low tech aids. Source: Pinterest.

4.2.2. Interviews

The following findings are drawn from a set of semi-structured interviews conducted as part of a broader interdisciplinary research effort involving a psychologist, a computer science researcher, and the author of this thesis. Two interviews were conducted directly by the author: one with the mother of a teenager with autism who had previously used an AAC system, and one with a speech-language pathologist holding a Master's degree in neurorehabilitation. Both interviews were carried out remotely with participants based in Colombia. In both sessions, the existing version of the AACreative application was demonstrated via screen sharing, allowing participants to see its functionalities and provide contextual feedback informed by direct observation of the tool. The remaining interviews were conducted by the psychology researcher within the same team, and involved speech-language pathologists, special education teachers, neuropsychiatrists, and AAC specialists working across clinical, school, and community settings in Italy. Findings are presented thematically and interpreted through a design lens, focusing on patterns relevant to the development of a tangible AAC storytelling tool.

The caregiver as an indispensable mediator

Across all interviews, the role of the adult, whether therapist, teacher, or family member, emerged as central to any successful AAC intervention. Professionals consistently described their work as triangulated between the child, the symbol system, and the people surrounding the child. As one speech-language pathologist noted, caregiver training is not optional but a structural component of the intervention: without an adult who understands, models, and sustains the system across contexts, even well-designed tools fail to take hold. This was confirmed from the family perspective as well: the mother interviewed described how the school inclusion teacher was decisive in her son's communicative development, and how the system only functioned meaningfully when it was embedded in a consistent relational environment. The implication for design is significant: any tool aimed at minimally verbal children must be conceived not as a child-facing interface alone, but as a cooperative system that actively supports the mediating adult.

Personalization as a condition of use, not a feature

No standard symbol set was reported to work equally across users. Professionals across all interviews emphasized that personalization is not an enhancement but a prerequisite: without it, children, particularly those with autism or significant cognitive delay, fail to recognize or engage with the symbols. "With autistic children or very young ones, you almost always have to photograph their own objects against a white background", noted the speech-language pathologist interviewed by the author. Several clinicians described starting from real photographs of the child's family, toys, and familiar places before gradually introducing more abstract pictographic representations. The mother interviewed confirmed this pattern from lived experience, describing how her son, a highly visual learner, did not respond to repetition-based methods until image-word associations were introduced in a way that matched his **perceptual profile**. The creation of personalized material was universally described as time-consuming and resource-intensive, pointing to a genuine unmet need for tools that lower the production burden while preserving contextual relevance.

Personalization was also described as extending beyond the visual dimension of symbols to encompass linguistic complexity. Several professionals noted that the grammatical and syntactic structure of AAC content must be calibrated to each child's developmental level, from single-word outputs for children at an early communicative stage, to multi-symbol utterances and simple sentence structures for those with more developed receptive and expressive skills. "We don't expect a child to produce a sentence if they are at the word level, both verbally and in AAC," observed one of the Carlo Besta speech-language pathologists. An AAC specialist working in schools further described the importance of controlling morphological elements such as articles and prepositions, adjusting their presence or absence depending on what the child is cognitively able to process. This dimension of personalization, linguistic rather than visual, has direct implications for AI-generated content: a system that produces narratives at a fixed complexity level risks being inaccessible to children at either end of the communicative spectrum, reinforcing the need for configurable output parameters that reflect **individual therapeutic objectives**.

Motor access as a defining design constraint

Several professionals emphasized that the physical dimension of AAC interaction cannot be treated as secondary to its linguistic or symbolic design. The speech-language pathologist interviewed by the author described a standard clinical practice in which an occupational therapist reviews each case before introducing a new tool, assessing whether external supports, such as armrests or adapted grips, are needed, and evaluating whether the physical demands of a new interaction modality might introduce motor fatigue. This collaborative assessment reflects a broader clinical awareness that the usability of any AAC system is inseparable from the user's **motor profile**. Conditions such as cerebral palsy, dystonia, and neuromuscular disorders significantly affect a child's ability to isolate finger movements, sustain pointing gestures, or navigate a touchscreen with sufficient precision and speed. As one of the Carlo Besta speech-language pathologists noted, "the ability to singularize movement is necessary" for touchscreen interaction, and for children who cannot do so reliably, the system must either offer alternative access methods or risk becoming inaccessible altogether. These observations underscore the importance of considering physical interaction not as a usability concern but as a clinical variable that shapes whether a tool can be used at all.

Narrative as a higher-order therapeutic objective

Storytelling was consistently positioned by professionals as a **more complex communicative function than requesting** or labeling, one that requires the integration of vocabulary, sequencing, and pragmatic awareness. "The story level, even a simple one, is a higher level. It requires the integration of word, sentence, and narrative," observed one of the speech-language pathologists at the Istituto Neurologico Carlo Besta. Despite this complexity, narrative was described as a valuable therapeutic tool precisely because it provides a motivating and socially shared context for practicing communicative skills. Several clinicians used short stories with pictogram support to work on sentence construction, vocabulary generalization, and comprehension. One therapist noted that "the story is accepted better than AAC because it is a shared experience", a moment of joint attention between child and caregiver that **carries relational**, **not just linguistic, value.**

The lack of adapted material as a structural barrier

Multiple informants described the production of personalized AAC material as one of the most demanding and underestimated aspects of their work. "The critical issue is that we don't have much material," stated one of the Carlo Besta logopedists directly. Also a special education teacher described spending significant time drawing and adapting content by hand. An AAC specialist working in schools noted that she still relies primarily on paper-based systems because producing digital content at the required level of personalization is too time-consuming. The consensus across interviews was that generative AI represents a meaningful opportunity in this space, not as a replacement for clinical judgment, but as a way to reduce the production burden and make personalized narrative material more accessible to caregivers and educators working with limited resources.

The resourcefulness that professionals and families display in the absence of adequate tools further illustrates the depth of this gap. The Colombian speech-language pathologist interviewed described the case of a mother who prepared PowerPoint presentations to narratively anticipate upcoming events for her autistic child, a self-made solution to a need that no available tool was addressing. This resonates with the experience of the mother interviewed by the author, who recounted using a picture book about a child visiting the dentist to help her son anticipate and emotionally prepare for the procedure. Both cases point to the same underlying need for these **social stories**[19]: narrative material that is personalized, **anticipatory**, and grounded in the child's real-life context, produced through considerable parental effort in the absence of more accessible alternatives.

Generalization as a persistent challenge

A recurring concern across interviews was the difficulty of transferring AAC skills beyond the therapeutic setting. Children who learn to use symbols in therapy often fail to apply them at home or school, a phenomenon that professionals attributed to insufficient training of non-specialist adults, inconsistency across environments, and tools that are perceived as too complex or too rigid to adapt. "We often find that children have a tablet with AAC but don't use it, not even in contexts where it would be useful," observed one speech-language pathologist. An educator working in schools described this as a daily challenge: "The autonomy children acquire at school is typically not generalized at home."

This pattern suggests that the value of any AAC tool is inseparable from its adoptability across contexts, and that design decisions around **simplicity, portability, and caregiver accessibility** are not peripheral concerns but core to whether the tool can be actually used.

Family resistance as a barrier to adoption

A less expected but consistent finding across the clinical interviews was the resistance of families to AAC systems, particularly among parents who fear that the introduction of alternative communication may delay or inhibit verbal language development. "Many parents see AAC as an obstacle to proper language development, they think that if children use these tools, they won't learn how to speak. But that's not true," stated a group of neuropsychiatrists and clinicians, whose observation was met with broad agreement from other participants. Professionals described this resistance as one of the most persistent challenges in their practice, compounded by the fact that many families feel they already understand their child's needs intuitively and see **no added value in a formal system**. This finding resonates with the experience of the mother interviewed, who described developing an informal family communication system, based on onomatopoeia, gesture, and shared context, that was effective within the household but collapsed when her son entered school and encountered adults who could not interpret his signals.

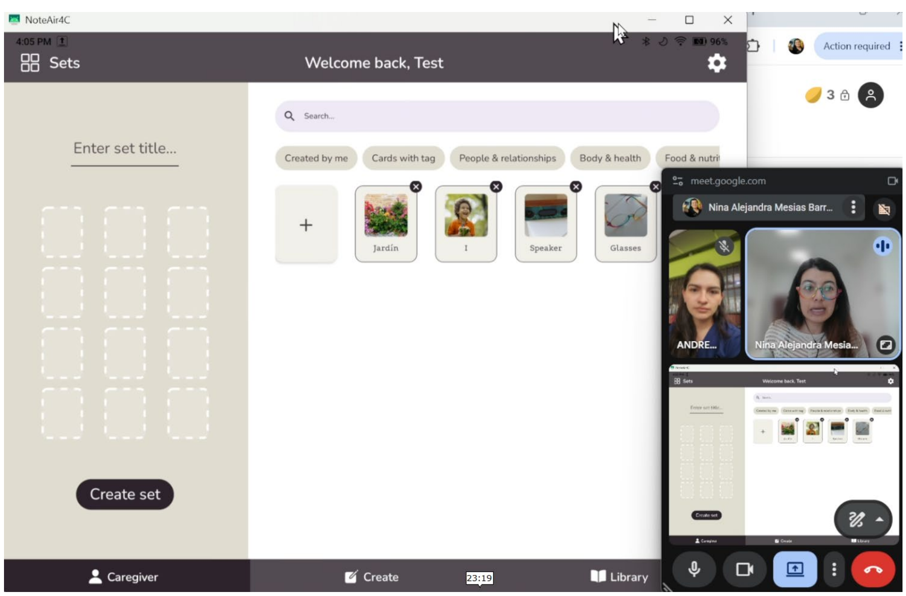


Figure 19. Remote Interview with a speech-language pathologist

4.2.3. Shared therapy sessions

The following activities were identified through non-participant observation of video-recorded therapy sessions shared by speech-language pathologists at the Istituto Neurologico Carlo Besta, complemented by their own accounts of clinical practice. The sessions involved a minimally verbal user who communicates primarily through Italian Sign Language (LIS) and were conducted in the presence of her caregivers. Two narrative activities were identified as particularly relevant:

Activity 1 — Storytelling from graphic symbol sequences

The therapist supports the user in telling a story using a sequence of four pictogram cards, the user first has to organize them in a logical order and then tell the story using in this case Italian sign language. Despite its brevity, the narrative contains the fundamental elements of a story: characters (a mother and a child), a setting (the street), objects (a car, a hat), emotional states (the mother is sad), and actions (walking, flying, stopping). The therapist asks the user to retell the story several times, slightly reordering the cards and prompting with questions ("why?", "who?") to encourage more detailed narration. The caregivers are then invited to retell the story themselves using sign language while repeating the words aloud.

What is particularly significant in this moment is the reversal of roles: the user begins to assist her caregivers by pointing to the graphic cards, producing onomatopoeia and even some words, and using subtle gestures such as nodding or shaking her head. She also actively encourages one of her caregivers to participate, expressing visible interest in having everyone repeat the activity. This observation highlights the role of graphic symbols not merely as communication aids, but as shared narrative anchors that structure turn-taking and invite active participation across all profiles involved in the session, including the child herself as a narrator and guide.

Activity 2 — Storytelling through pictogram dice

The therapist begins by presenting the faces of six dices (Figure 20), each side displaying a **pictogram**, and explains their meaning through sign language. She also asks the user questions about the words each pictogram might represent, activating prior knowledge before the activity begins. Each dice corresponds to a different word category, generating varied narrative possibilities depending on which faces appear after rolling.

The therapist selects four of the six rolled dice and constructs a brief story by weaving together the pictograms that appear, narrating it in sign language while repeating the words aloud. One of the caregivers then rolls the dice and invents their own story from the resulting images. Finally, the user repeats the activity independently.

Three design-relevant observations emerge from this activity. First, the element of choice: the user selects which dice to incorporate; introduces a sense of authorship and agency that visibly increases engagement. Second, the user encounters notable difficulty when attempting to add more words to construct a fuller narrative, suggesting that the linguistic demand of story composition represents a significant barrier for minimally verbal children. Third, besides choosing the user organizes the dices before telling the story. These observations raise attention on the need of **sequencing** for storytelling, and questions about how a tangible interaction system, combined with a narrative support tool, could lower the expressive threshold without replacing the child's communicative initiative.

Figure 20. Example of dices used for story generation in the context of speech therapy.

5. DESIGN PROCESS 

5.1. VERSION 1

Based on the information collected, the first physical sketch was built to allow experimenting with an initial interaction flow proposal from the perspective of both profiles: the caregiver and the AAC user. The Figure 21 shows a draft of the first idea.

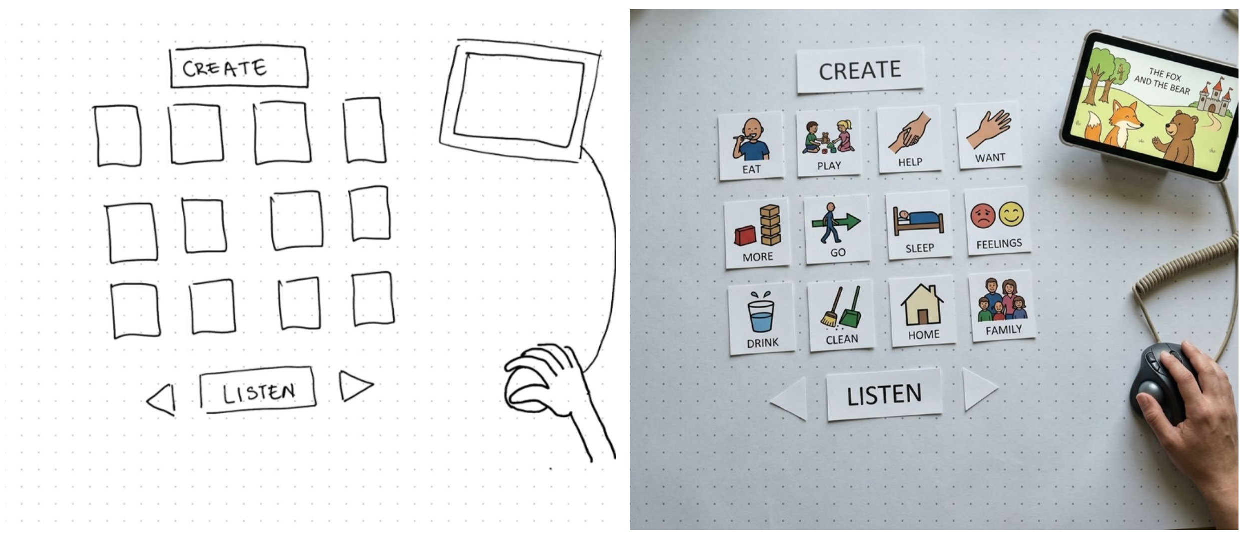


Figure 21. Ideation process to generate a first design concept (right image generated with Gemini AI)

For its development, an ESP32 board was used along with an RFID reader and a vibration motor for giving haptic feedback. The ESP32 board was connected via serial cable to an Android tablet running an application that simulated the functionalities of the AACreative application, without having real access to the symbol database, but instead using a local bank of ARASAAC[3] symbols and a previously defined story. This app was developed using Flutter. Although UDP communication was initially tested to avoid the use of a cable, this approach was discarded because WiFi communication proved insufficiently stable. The Figure 22 shows the components of physical sketch 1.

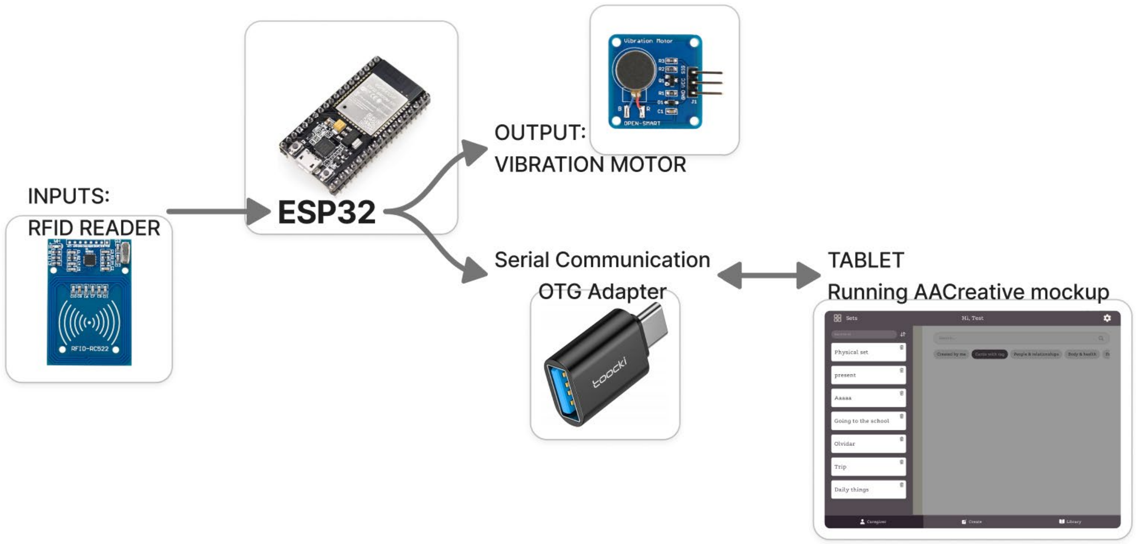


Figure 22. Hardware involved in physical sketch 1.

The application had the same three screens as the real app:

Caregiver: This screen showed the chip: card with tags and contained only the button add (Figure 23) to associate physical cards and generate a digital representation of them. Figure 24 shows the overlay that appeared once the NFC tag was read.

Create: On this screen, the cards selected through physical interaction appeared and disappeared if the same NFC tag was scanned again (See Figure 25).

Library: Here the images containing the story were displayed, and an audio track narrating the content of the visible page could be activated.

The exchange of information between the app and the ESP32 was based on sending either the number of the scanned tag or a control command. In this version, the tags corresponding to the cards were stored locally in the app to identify repeated readings and remove them when that was the case.

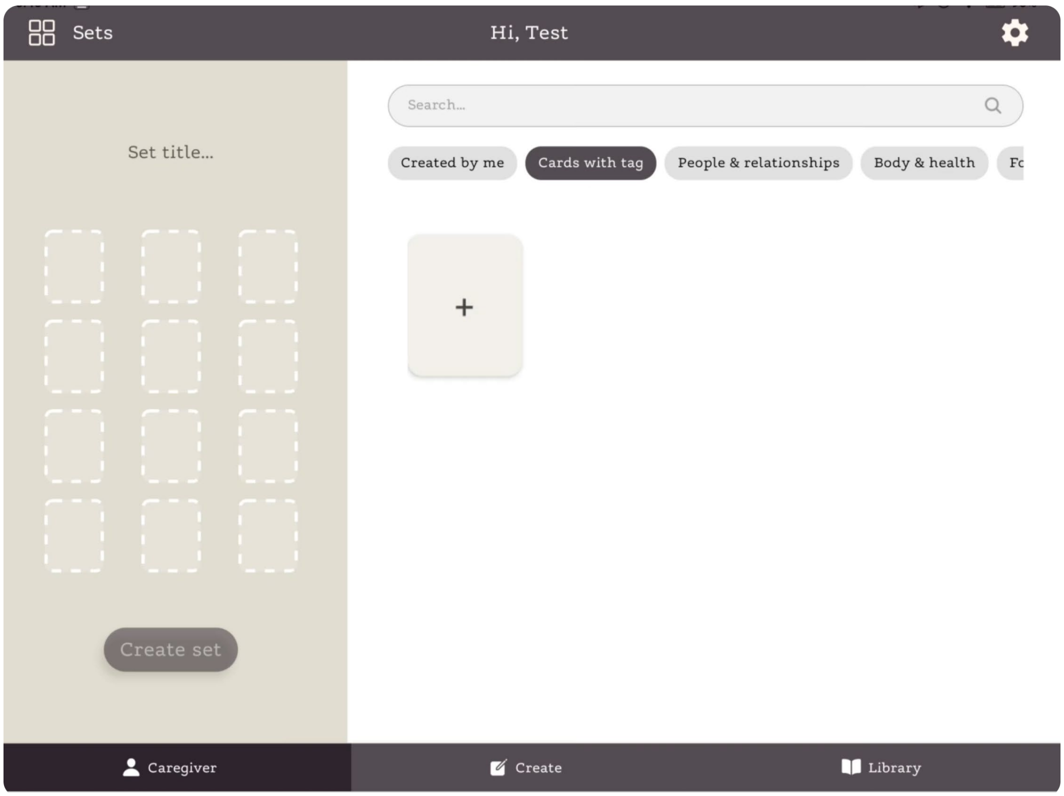


Figure 23. Screen for the caregiver, to associate the physical cards with their digital representation.

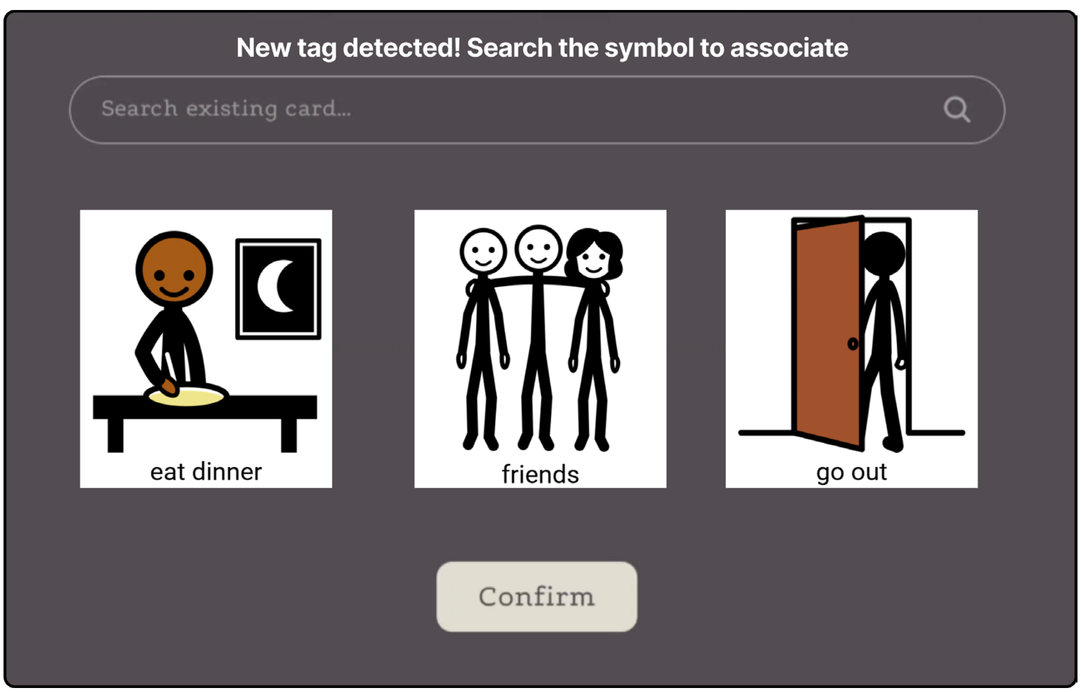


Figure 24. Screen to select the digital representation of the last NFC tag read. Overlay of the caregiver section

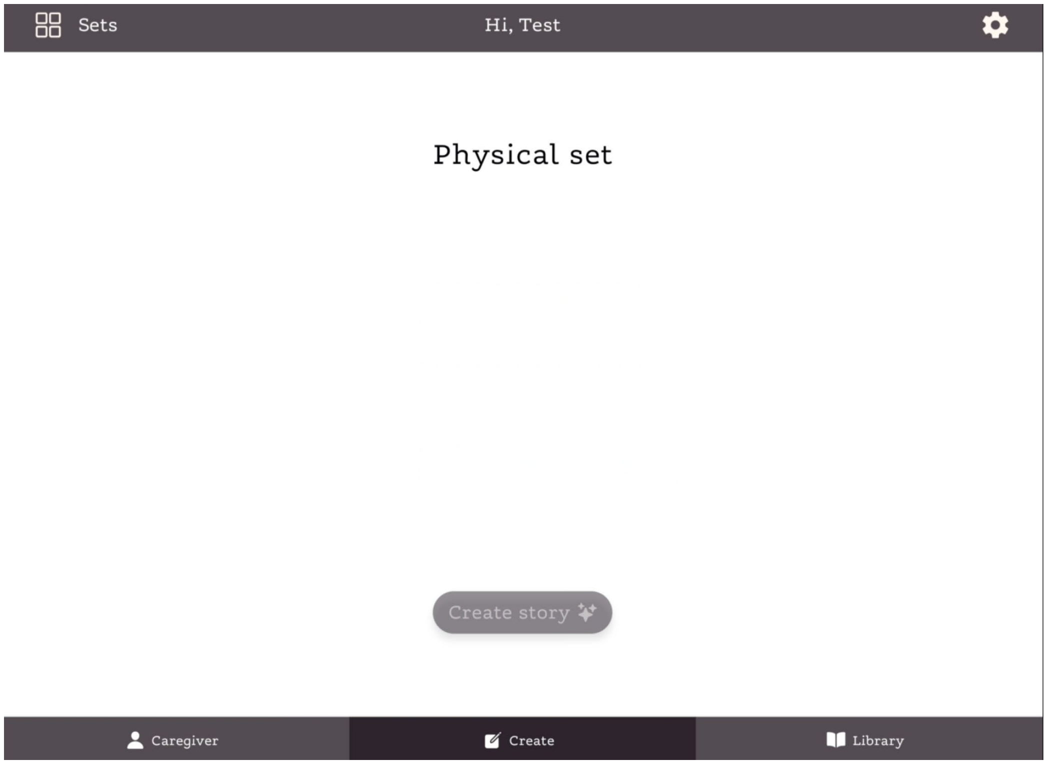


Figure 25. Screen for the interaction of the AAC user.

The physical setup consisted of a board arranged in a three-by-three grid, with nine square slots of approximately 8.5 cm per side, each designed to hold a physical NFC card. A separate, smaller board was placed in an additional row below the main grid, containing four control buttons: Create, Listen, Next Page, and Previous Page. Figure 26 and Figure 27 show the arrangement of all the elements of the prototype and the detail of the control board for scanning the cards.

Figure 26. First design concept.

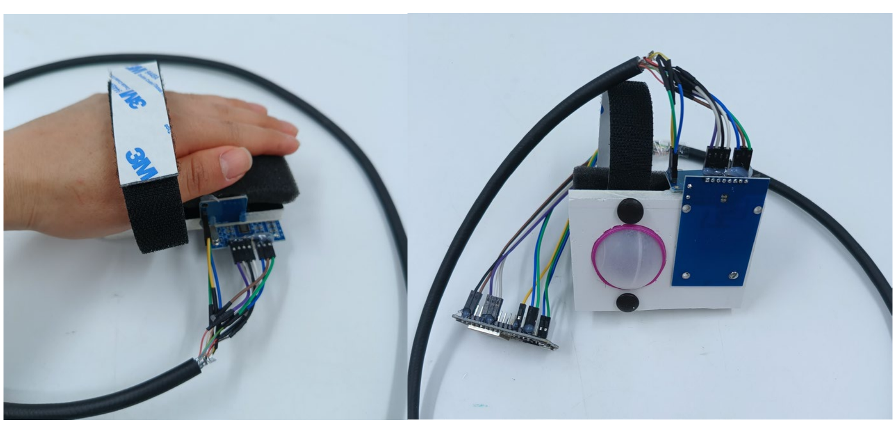


Figure 27. Controller of the first design concept

This prototype was presented as a discussion element in a group session attended by three researchers from the i3lab laboratory and the laboratory director (Professor Franca Garzotto). The design department faculty member supervising this thesis also participated (Professor Daniela Petrelli). Subsequently, a peer review session was conducted with another student from digital and interaction design with a background in product design. From these experiences, conclusions were drawn for the development of a second physical sketch:

The use of a serial cable to connect the controller to the tablet was identified as a significant constraint, as it limited the freedom of movement across the board during interaction. The spherical form of the controller also generated unexpected interaction patterns: participants instinctively attempted to insert it into the holes of the board, treating it as the contact point for scanning, which revealed that its shape functioned as an unintended affordance, however the sphere didn't fit in the holes of the control board. Further concern was raised around the two-step process of creating a custom card and associating it with a physical tag, which participants experienced as disconnected rather than as a single unified action. Finally, the sets overlay was noted as requiring manual dismissal after a selection was made, a small but recurring friction point in the interaction flow.

After this session an agreement was reached to organize an encounter with the experts at the Istituto Neurologico Carlo Besta to show an improved version of both the application and the physical sketch.

5.2. VERSION 2

The second version of the skecth in hardware was designed to integrate the physical interaction layer into the real AACreative application. To achieve this, access was granted to the application's repository and the ESP32 firmware generated in the previous version was adjusted accordingly. In this version, Bluetooth Low Energy communication was used between the tablet and the ESP32, which required the addition of a battery to power the

ESP32 board. It was necessary to reconnect all components and reorganize their physical arrangement to make the physical sketch manageable (Figure 28).

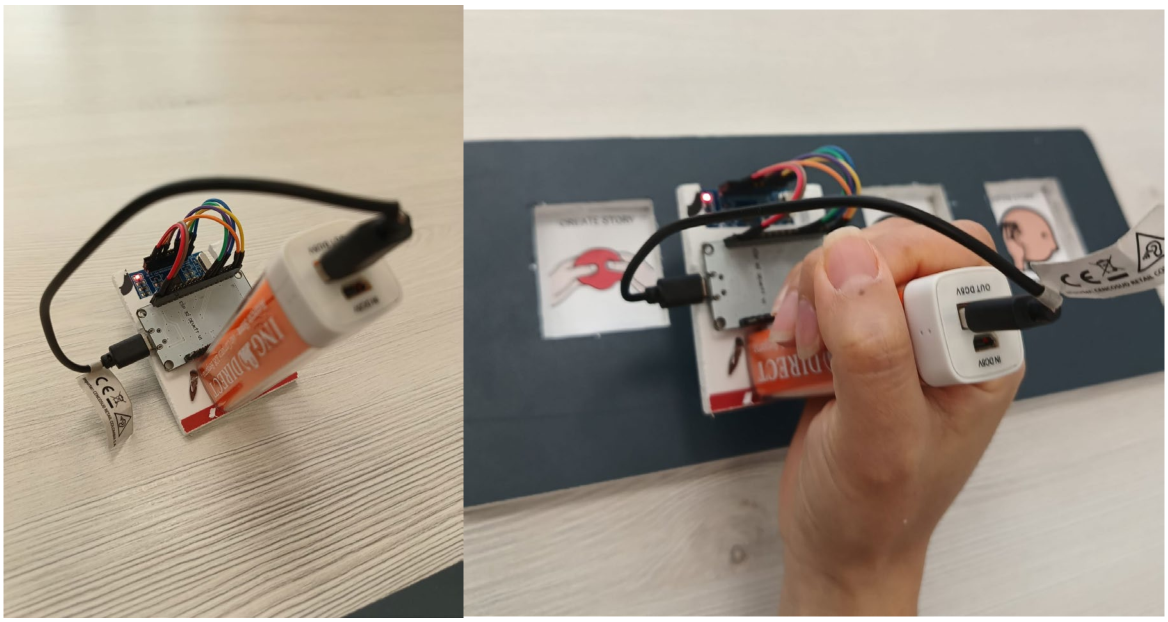


Figure 28. Controller of prototype 2. Since in this version the ESP32 is not connected to the tablet a battery is needed.

The operational flow in this version is as follows:

Caregiver side

1. The caregiver taps the add button within the "cards with tag" chip in the Caregiver section (Figure 29). This establishes the Bluetooth connection between the two devices in the first tap.

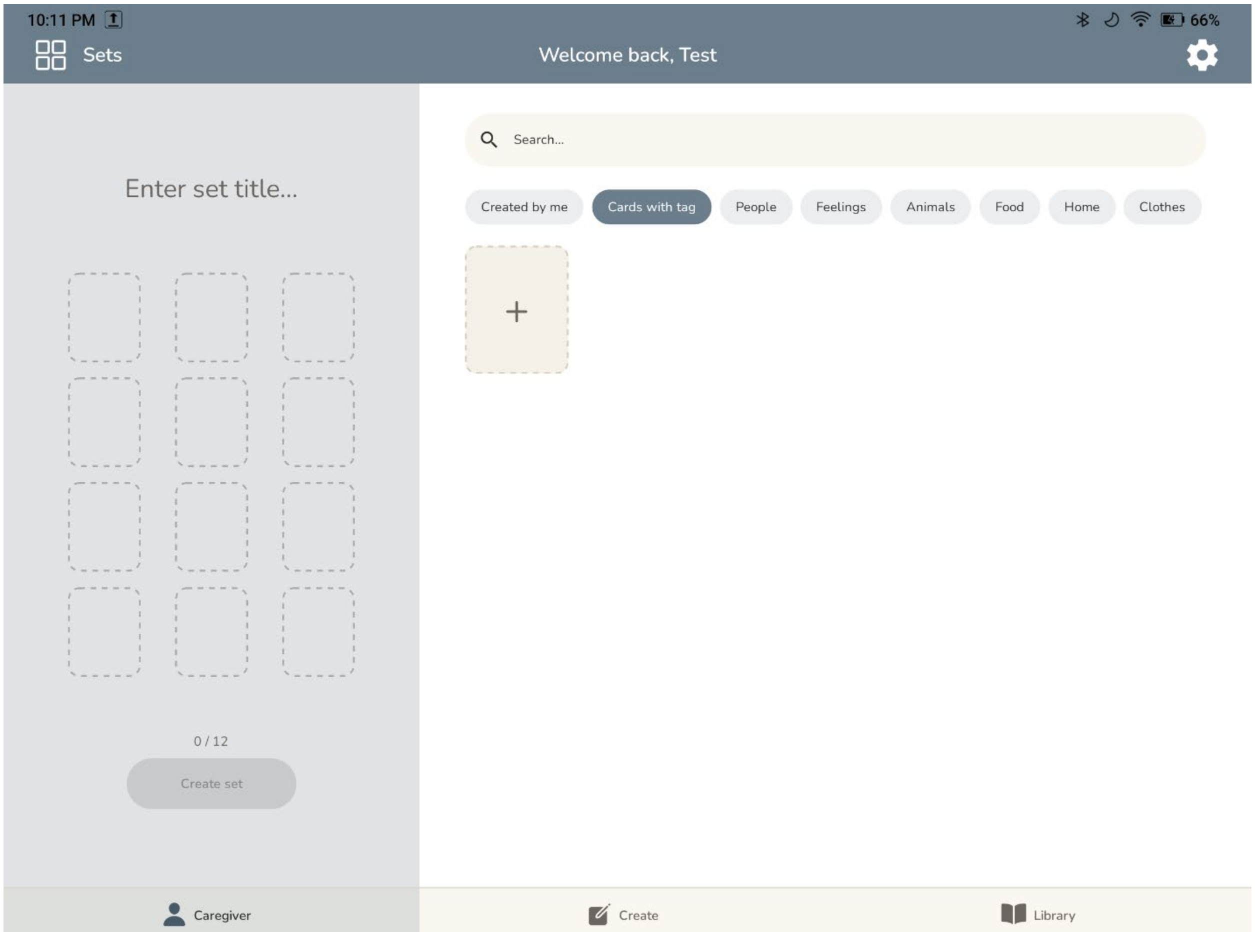


Figure 29. Caregiver section where the "Cards with tag" filter appears to create an association with physical cards

2. An overlay opens, waiting for an NFC tag to be detected (Figure 30).When a tag is read, the overlay closes and a second overlay appears, allowing the caregiver to

search for the corresponding digital card. Custom cards are loaded by default to speed up the association process (Figure 31).

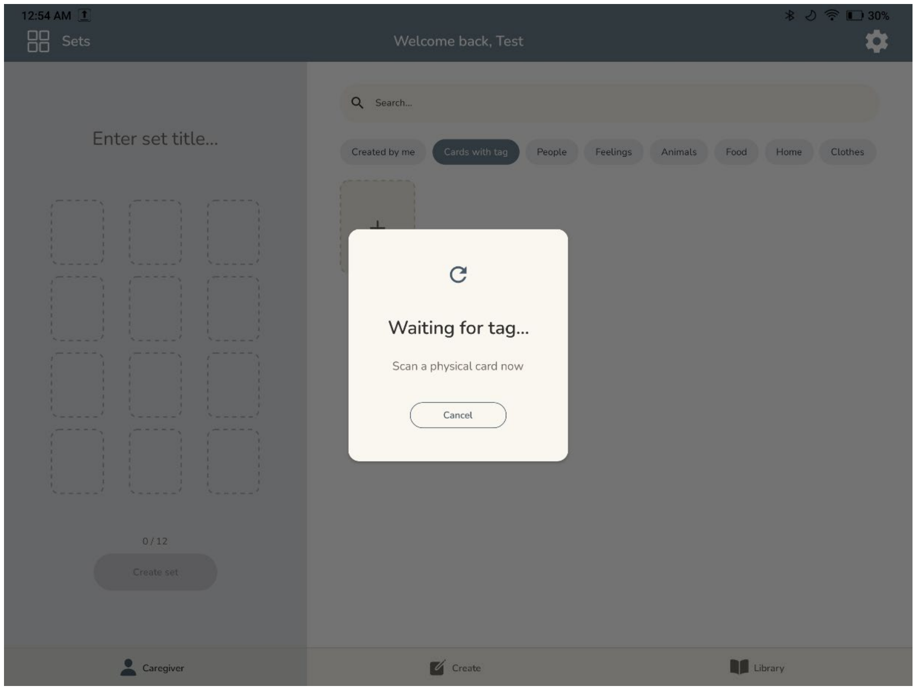


Figure 30. Overlay for waiting the scanning process of a physical card.

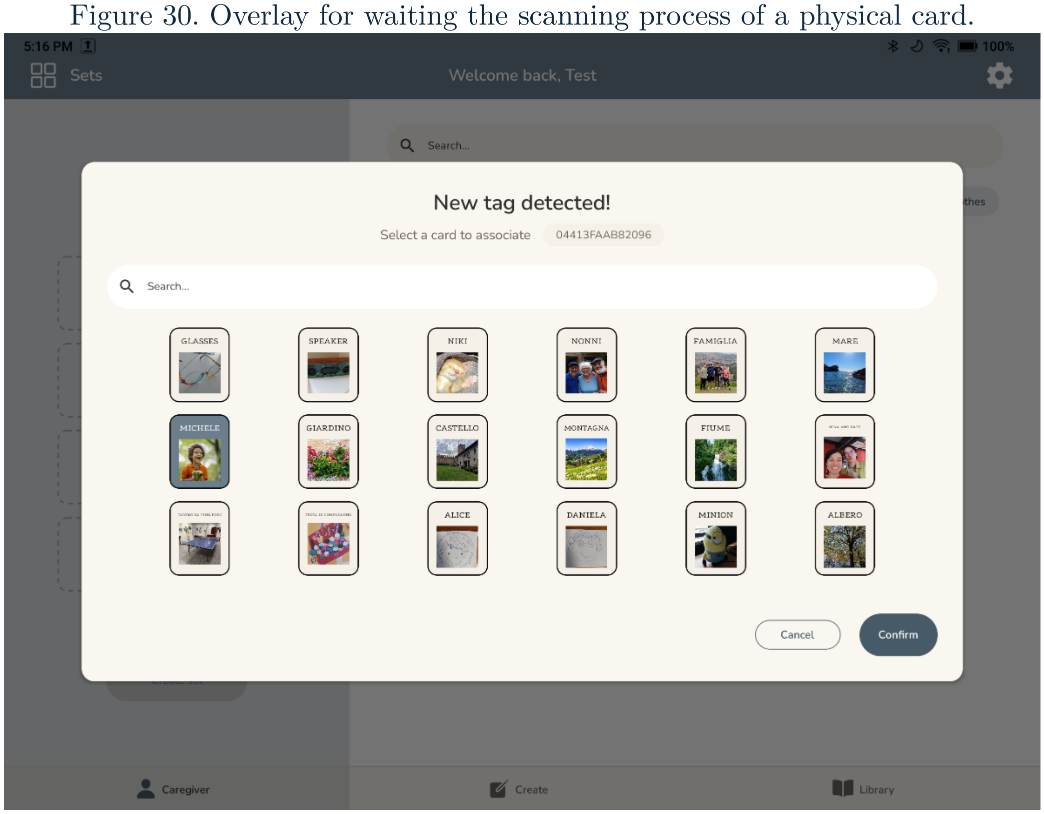


Figure 31. Overlay to associate the physical card with its digital version.

3. If the digital equivalent of the physical card does not appear among the custom cards, the search bar can be used to explore the ARASAAC database.
4. The corresponding digital card is selected.
5. The caregiver taps confirm.
6. Steps 1 through 6 are repeated for each physical card.
7. The caregiver taps the sets button.
8. From the sets overlay, the first item, corresponding to the physical set, is selected. This option is only visible when a Bluetooth connection is active (Figure 32).

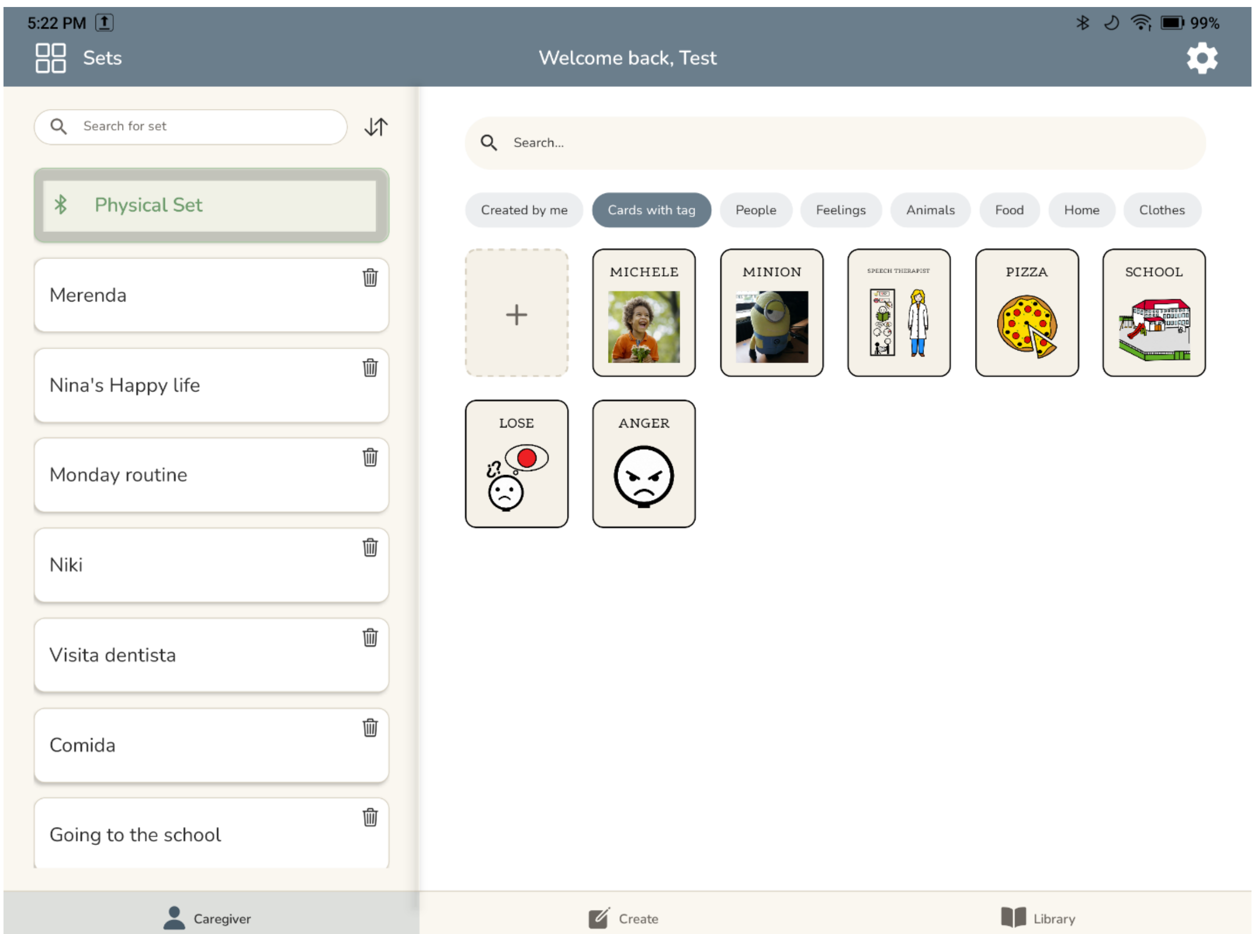


Figure 32. The Physical Set appears as the first option in the Sets list if there is a Bluetooth connection stablished.

Child side

1. The physical set screen appears empty and the system waits for one of the previously paired cards to be scanned.
2. When a card is physically scanned, it appears on screen and its label is read aloud (Figure 33).

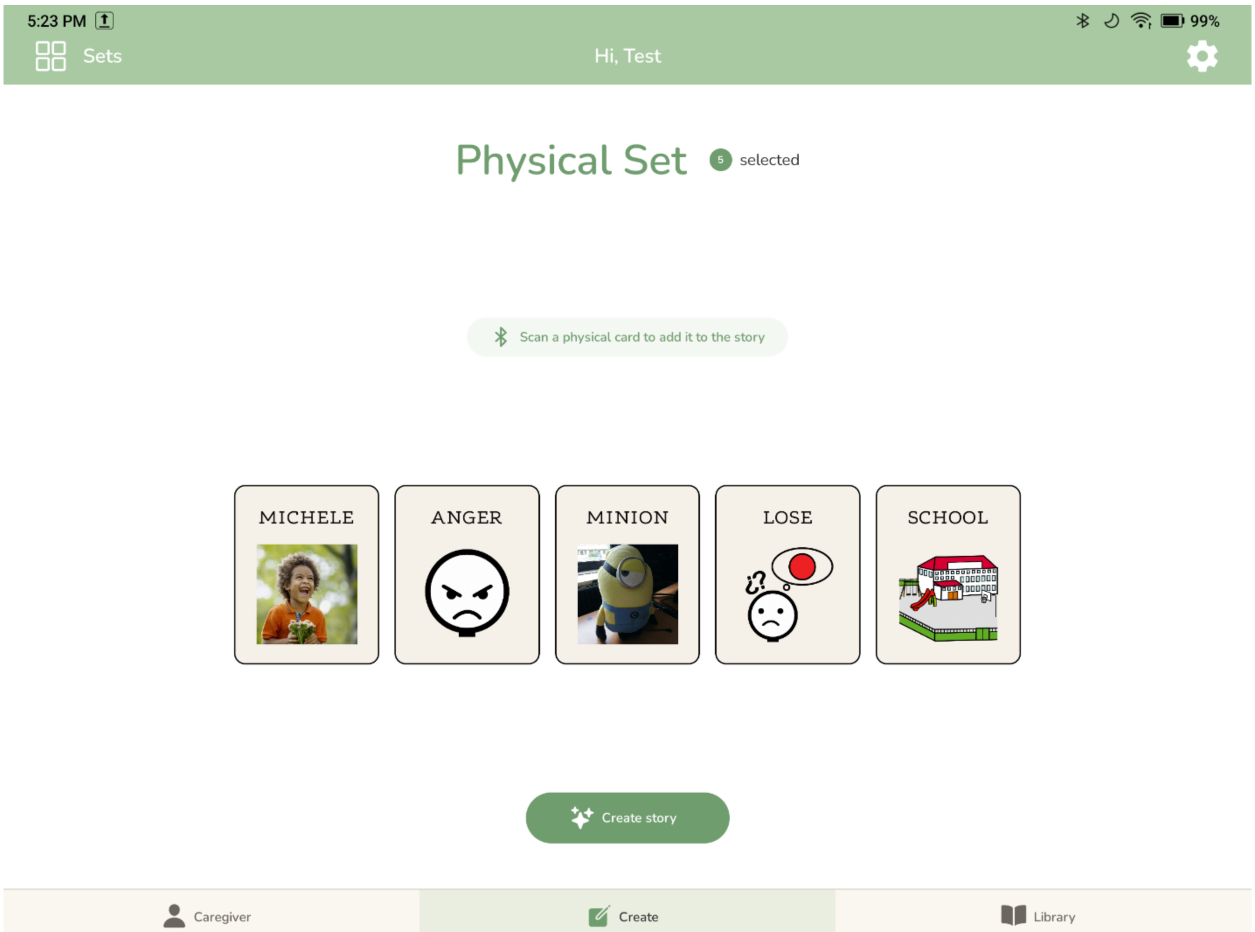


Figure 33. In the Create section, when a physical card is scanned it appears on the screen and is read aloud.

3. (Optional) If a card already displayed on screen is scanned again, its audio plays once more, the card is removed from the screen, and the user receives haptic feedback from the controller.
4. The step 2 is repeated per each card that will become part of the story
5. When the *Create card* is scanned, a control command is sent to the tablet to initiate the story generation process.
6. The screen updates to display a loading indicator while the story is being generated (Figure 34).

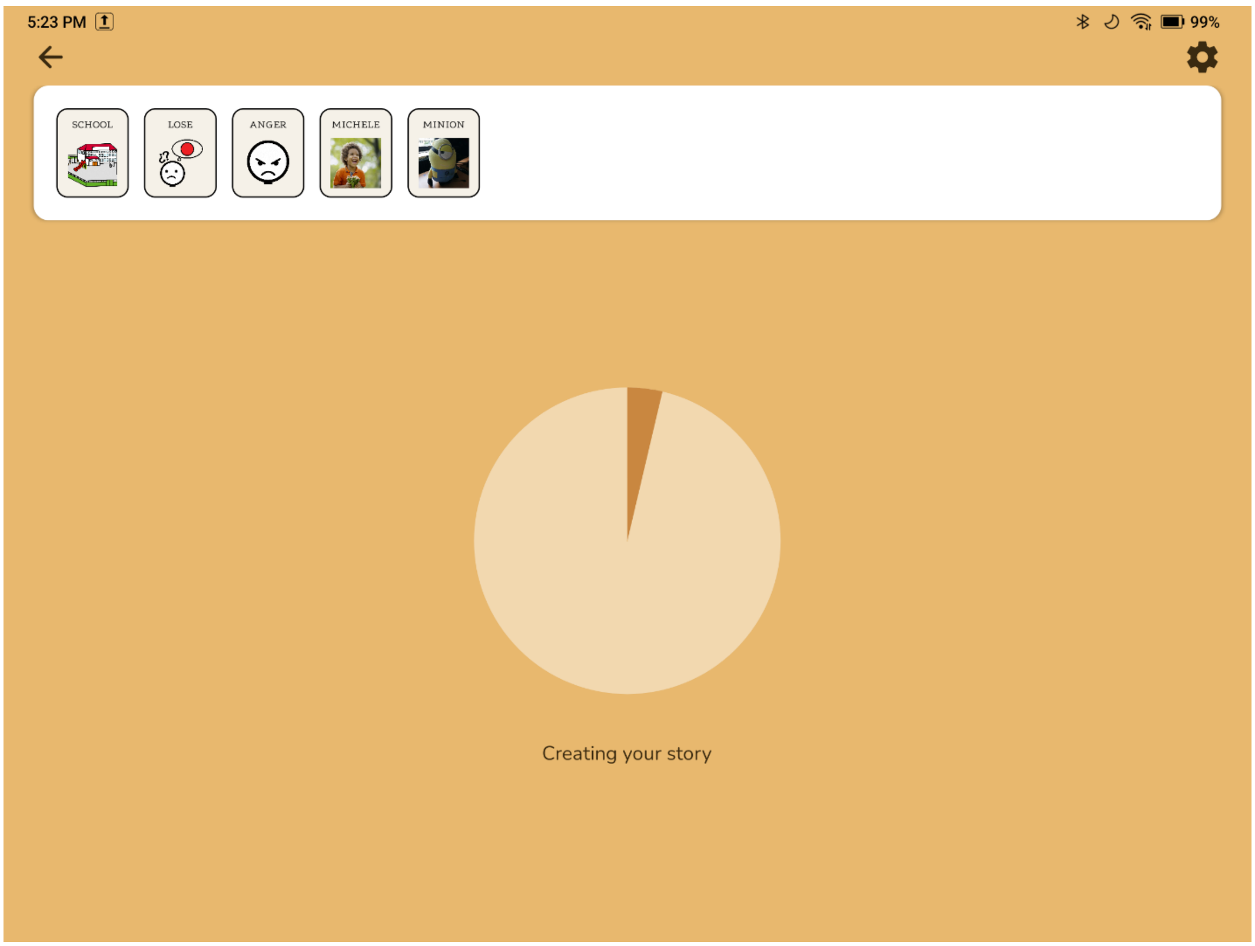


Figure 34. Once the "Create story" button has been scanned the creation process starts and an animation is shown,

7. Once the story is ready, the first page is displayed and the system awaits further control commands. At this point, scanning cards that are not control commands has no effect on the system.
8. The *Listen card* is scanned and the current page is read aloud.
9. The *Next card* is scanned and the story advances to the following page.
10. This process is repeated for the remaining pages (see Figure 35). Optionally, the Previous card can be scanned to navigate the story in the opposite direction.

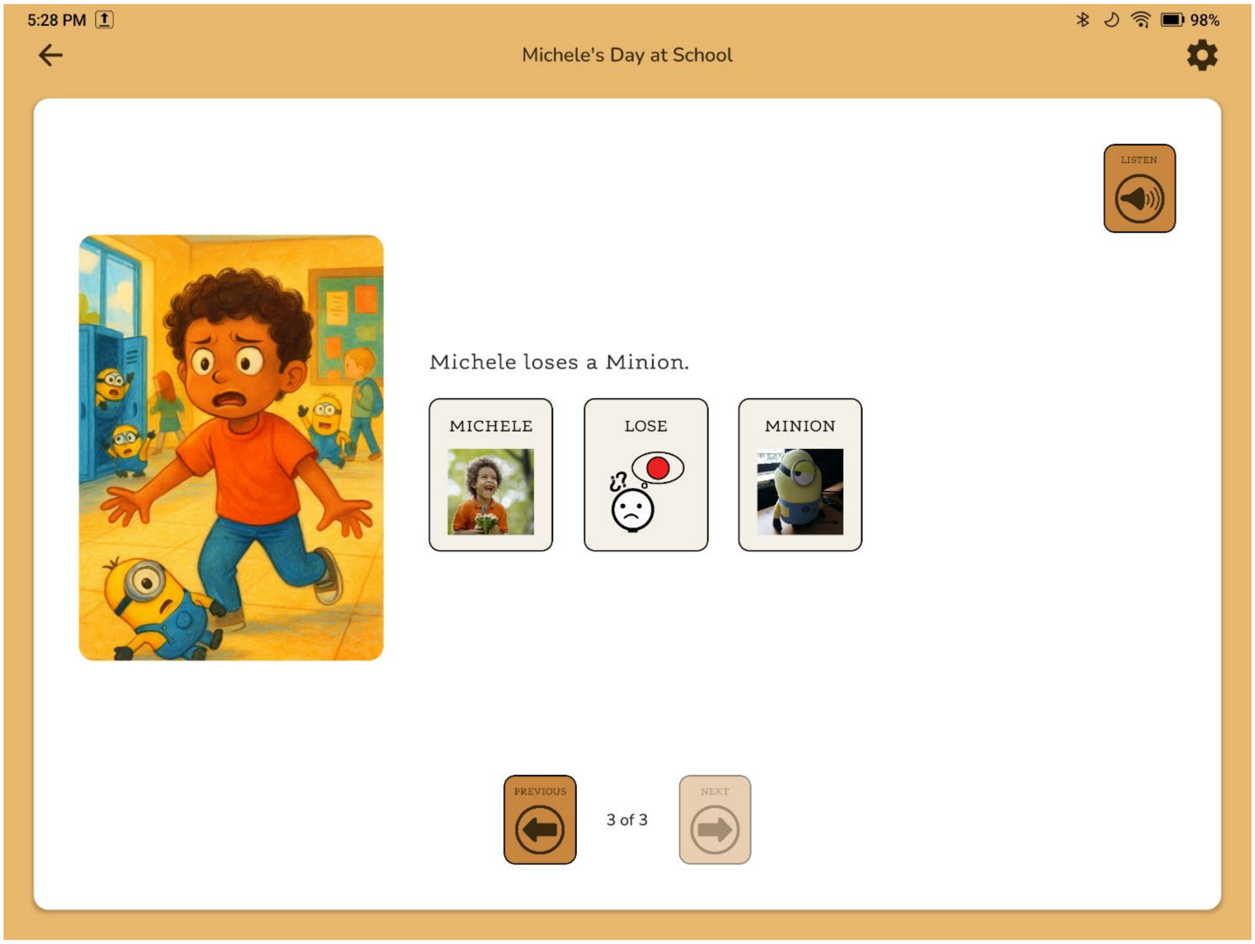


Figure 35. Story navigation can be done scanning the physical buttons.

The logic behind is shown in Figure 36, which explains the message exchange between the two devices depending on the phase of the process. Figure 37 and Figure 38 summarize the physical-digital interaction sequence for the caregiver and the AAC user respectively.

It is worth noting that the intended design keeps the AAC user's interaction entirely within the **physical layer**, with no need to interact with the tablet directly.

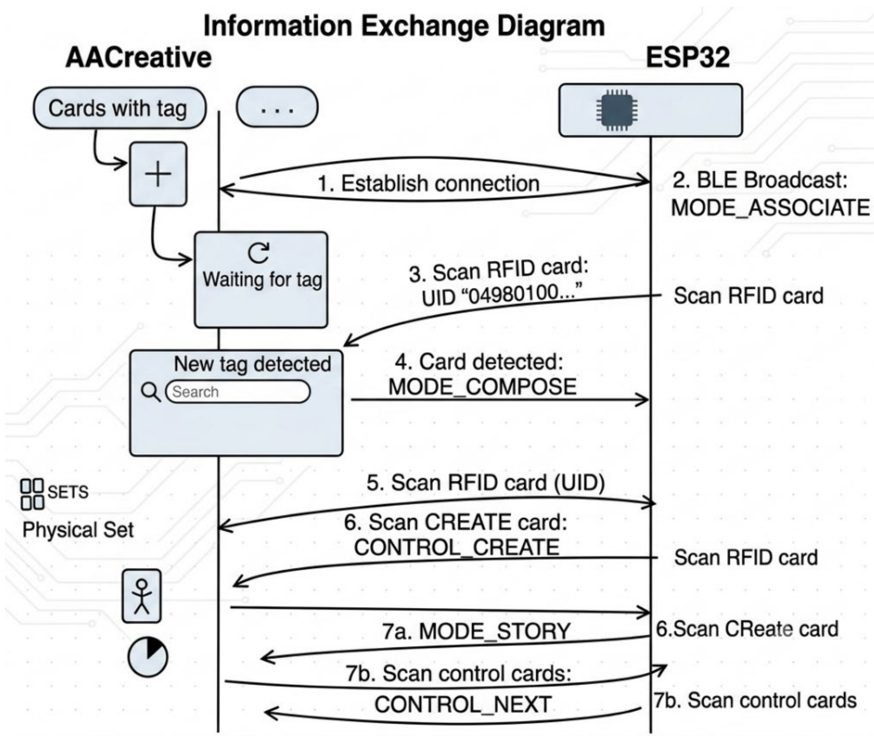


Figure 36. Communication between the AACreative App and the tangible controller. Image generated by Gemini AI based on a drawing by the author.

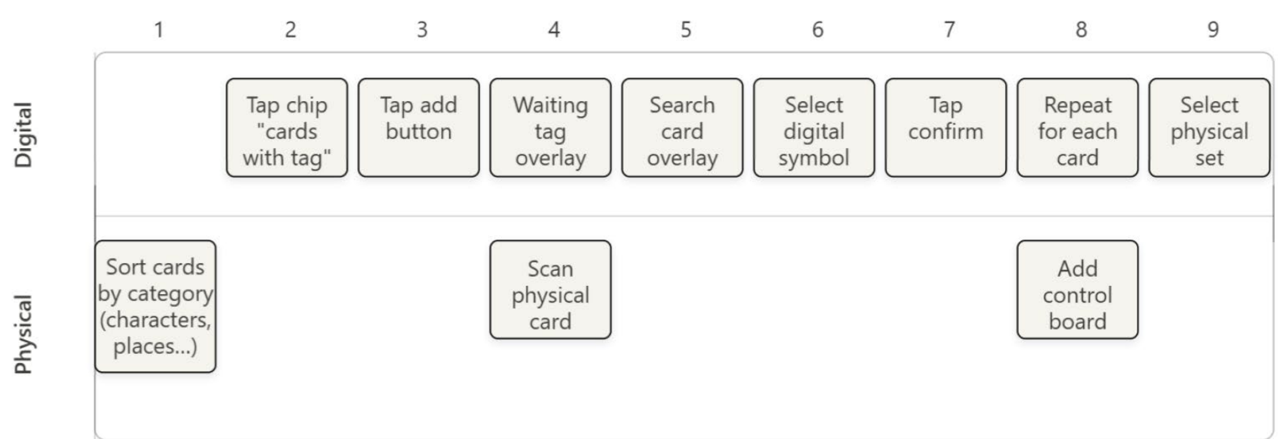


Figure 37. Physical-Digital Interaction Map for the caregiver

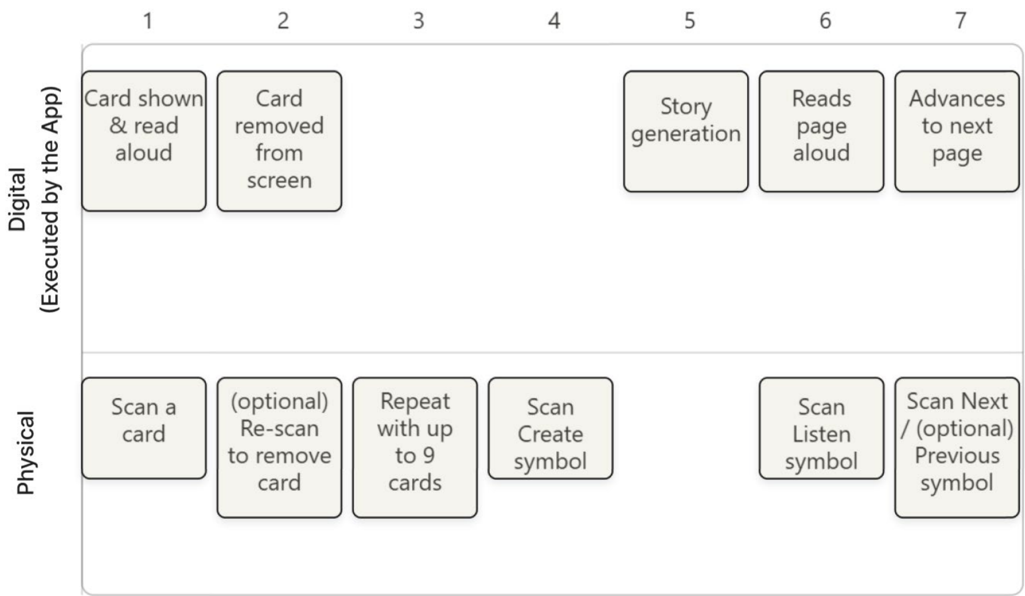


Figure 38. Physical-Digital Interaction Map for the AAC user

This version introduces a new way of holding the controller, removing the sphere proposed in the first version. The card organization board and the control board, on the other hand, remained unchanged.

5.3. CODESING WORKSHOP

In early June 2026, a co-creation session was held at the Istituto Neurologico Carlo Besta with a group of professionals working with minimally verbal children, all had relevant clinical experience with the population for whom AACreative is designed. The session was not conceived as a formal evaluation of the second physical sketch, but rather as a space for dialogue: the piece of hardware connected to the app served as a discussion tool to explore existing therapeutic dynamics around storytelling and to understand how a system of this kind might fit into real clinical practice. The insights gathered during the session informed subsequent design decisions for both the physical sketch and the AACreative application.

Preparation

After generating multiple stories in the AACreative application, the potential of custom cards became evident: the resulting stories incorporate them not only into the narrative but also into the illustrations, producing a greater sense of identification with the story; Figure 39 shows how the character in the story is very similar to the cat in the personalized card.

Figure 39. Example of the use of personalized cards in the storybook

For this reason, a set of custom cards with photographs presenting **characters and places** was prepared specifically for the workshop. To provide a wider range of narrative elements, additional cards were also prepared for the categories of **emotions, situations, and verbs**, using pictograms sourced from the ARASAAC portal.

Execution

The attending therapists, three speech-language pathologists and a psychologist specializing in child psychotherapy, were asked to engage in a role-playing exercise in which each participant represented a different profile: the caregiver, the therapist, and the

AAC user. The fourth therapist did not have a specific assigned role but participated actively throughout the session. The tasks assigned to each role were as follows:

1. **Caregiver:** Create a custom card by taking a photograph and uploading it to the system, then assign a name to the card.
2. **Therapist:** Select at least four physical cards, including the custom card created by the caregiver. Scan them one by one and complete the pairing process with their digital equivalents. Then arrange the cards on the board and select the physical set from the sets overlay in the Caregiver section of the application.
3. **AAC user:** Scan the physical cards of their choice from those arranged by the therapist, then use the control board to scan the cards for creating, listening to, and navigating the story.

Participant engagement was highly active and collaborative from the beginning. The group planned the story elements in considerable detail before starting, which led them to select nine cards and fill the board completely (Figure 40). Each participant carried out their assigned role, but the group maintained an ongoing dialogue throughout.

From the beginning, there was a strong desire to have control over the story. Participants discussed how the selected elements would come together to form a narrative and went so far as to imagine its development in detail. This dynamic prompted a broader discussion about how a tool such as the one proposed could be incorporated into a therapeutic session, raising concerns such as the potential frustration an AAC user might experience upon seeing a story that does not match their expectations. The resulting story, while it contained all the selected elements, lacked logical coherence and revealed limitations in the application's ability to construct narratives beyond subject-verb-complement structures.

Figure 40. Participation in the workshop: Preparation of the cards for the story creation

5.4. INSIGHTS

There was a strong interest in having greater **control over the story outcome** and the complexity of its elements depending on the developmental level of the AAC user. Personalization in this regard is particularly important given that each case has its own particularities. The therapeutic purpose of the tool was also discussed, and one of the therapists stated:

"It is not a communicative tool (…) the purpose of this activity is narrative and creative, in the sense that: I see what the story is like and I retell it. I choose the element and I enjoy the experience with the speech therapist to see what happens, we listen, and we see how it surprises us."

The group discussed which therapeutic activities could be associated with the tool. One possibility raised was:

"One activity could be: I know you went to a party, we tell it first with the images and then we see what artificial intelligence does with our images (…) That is a slightly higher level."

Another activity they imagined would consist of preparing a set of cards representing an event that happened to the AAC user, generating a story using those cards, and then having the user indicate whether the generated story resembles or not what they actually experienced.

Participants expected the **order in which cards are selected** to have implications for the resulting story, and the concept of **sequence** was highlighted as a relevant learning objective in the therapeutic context.

The therapists noted that the "when" element was missing from the story, since in their current practice they ask users to recall and narrate past events, asking for instance: “tell me what did you do yesterday?". In the same direction, they identified the core elements needed to compose a story as: who, where, what, and when. They also described the retelling activity they use in their sessions, which consists of:

1. The therapist tells a story
2. The child retells it and the accuracy of the retelling is evaluated
3. The user is then asked questions about the story: why, who, how, etc.

It was also mentioned that the tool could support more autonomous use, but the group emphasized in any case the importance of clarifying the therapeutic objective for each session. This could be reflected in the app as an option to include or exclude a comprehension check activity, such as placing story events in the correct order, which is one of the exercises the experts currently use.

Specific comments on the application

Participants noted that the pictograms selected as story inputs deserved greater visual prominence, both during the generation phase, where displaying them at a larger size could help maintain the user's focus while waiting, and within the resulting storybook itself. The waiting animation was identified as not enough engaging, with suggestions including a character-based animation depicting the story being prepared, or AAC symbols prompting the user to wait. On the caregiver side, the ability to rename custom cards and make meaningful use of the description field was flagged as a missing feature, as the current interface presents a usability issue that prevents the entered information from being saved. The sets overlay was also noted as requiring manual dismissal, adding unnecessary friction to the workflow.

Regarding the storybook itself, participants observed that the generated stories lacked an opening and a closing page, and that the narrative connection between pages was weak, with characters not being properly introduced before appearing in the action. Despite these limitations, the level of detail in the AI-generated illustrations was consistently highlighted as a strength. At a structural level, the subject-verb-complement sentence pattern that currently dominates the generated text was considered too limited to produce a fluid and engaging narrative, and the tool was seen as most appropriate for **children with milder impairments** in its current form. Two additional suggestions emerged: the introduction of color coding to categorize cards by grammatical role, and the guarantee that all symbols selected as input would be explicitly reflected in the resulting story.

Observations on physical interaction

The grid arrangement of the card frames prompted a suggestion to explore a single-row configuration instead, on the grounds that a linear sequence of cards more naturally conveys the concept of narrative order, an association that carries direct relevance in therapeutic contexts where sequencing is an explicit learning objective. The two-step process for creating and associating a custom physical card generated notable confusion: one participant declared that the card had already been created after completing the photograph, unaware that a separate association step with the physical tag was still required, pointing to the need for a more integrated flow. A more practical concern was also raised regarding the physical board itself, as cards were found to get stuck in the slots, making them difficult to remove during interaction.

On the feedback side, participants expressed a clear expectation of auditory or haptic confirmation once the Create story tag had been scanned, noting the absence of any signal at that critical moment in the interaction. The automatic narration of each page upon turning was proposed as a desirable default behavior, with the alternative of offering different control board configurations depending on the therapeutic objective. A reordering of the control board cards was also suggested, placing them in the sequence: Create, Listen, Next, and Back, to better reflect the natural flow of the activity. Finally, the idea of an additional board dedicated to comprehension assessment was introduced, through which the therapist could prompt the user with questions such as who, what, and where after the story had been read.

Finally, the session prompted a comparison between the application and the story dice the therapists use in their practice, noting that while the app replaces the narrative challenge for the AAC user, it could serve as a tool for learning how to tell a story and later adapting it. One of the experts also mentioned that she already uses AI in her therapy sessions to prepare materials, for example creating stories that contain a specific phoneme repeated multiple times, along with illustrations for coloring.

6. DESIGN ADJUSTMENTS PROPOSALS

6.1. SUGGESTIONS FOR REDESIGNING THE AACREATIVE USER EXPERIENCE

Based on the feedback gathered throughout the process and after repeated interaction with the AACreative application, a series of UX adjustments are proposed for each section of the interface.

Caregiver section

To simplify the pairing process between custom cards and the physical cards that represent them, a tag association button can be incorporated directly into the custom card creation flow (see Figure 41). This way, the caregiver does not perceive the process as two separate steps, and the overall task flow is simplified.

Additionally, to facilitate pictogram recognition and classification, color-coded borders are proposed for the cards. These borders can be toggled on or off in the application settings. The proposed system follows the color coding used by ARASAAC[33], which is based on a modification by McDonald and Schultz of the Fitzgerald Key. This system groups pictograms into six colors according to their grammatical role within a sentence (see Figure 42)[28]. This choice is further supported by the fact that this color system is used by the most widely adopted applications reviewed in the benchmarking section.

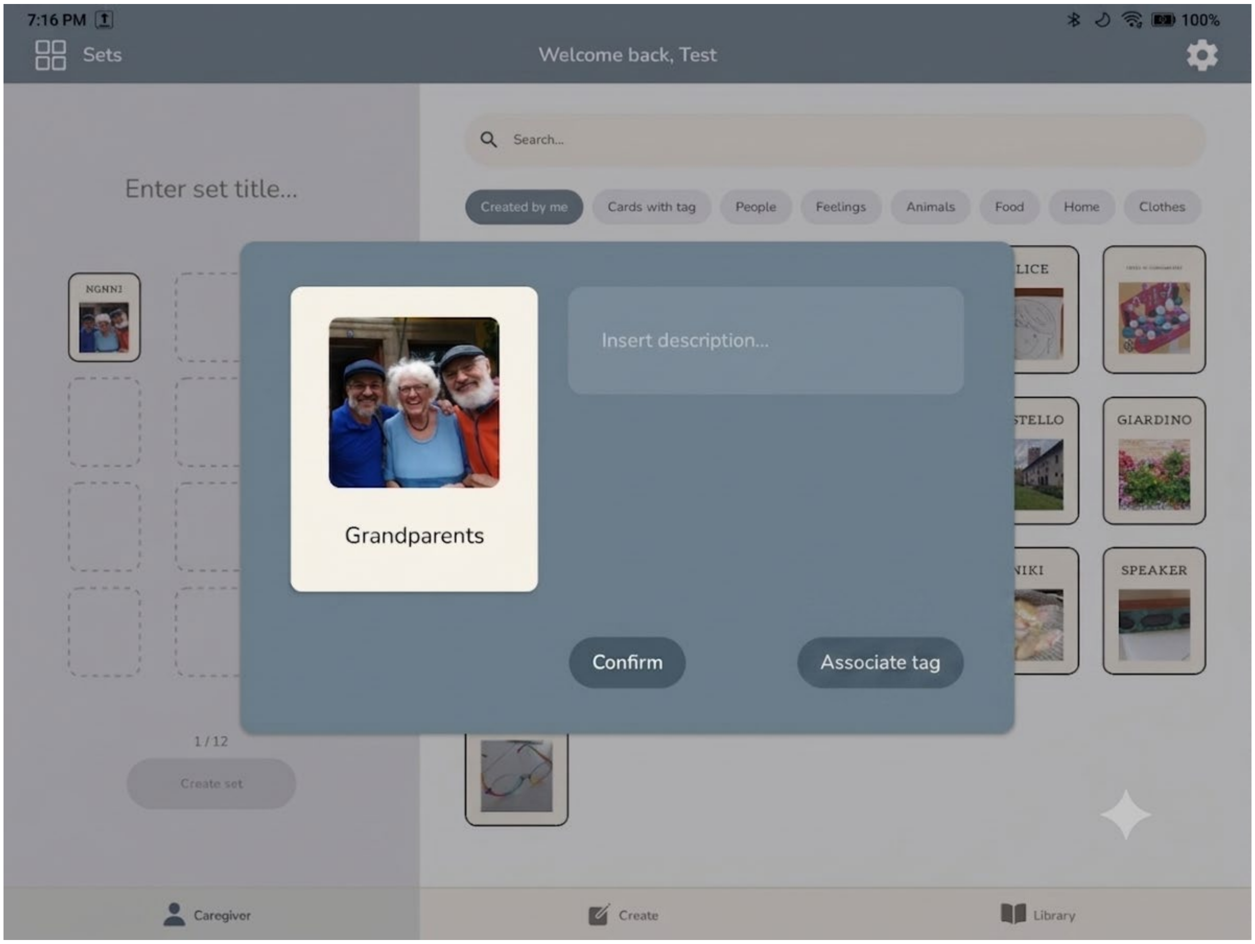


Figure 41. Addition of the button "Associate tag" in the section of personalized cards. (Made with Gemini AI)

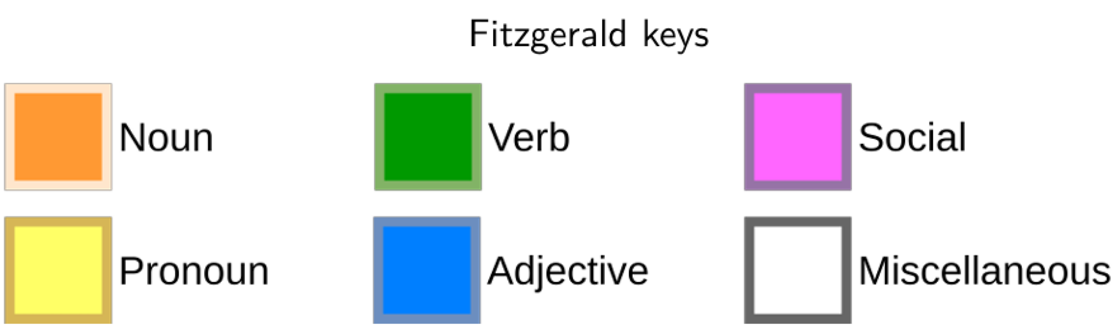


Figure 42. Colors used to categorize the words in the modified Fitzgerald keys system.[28]

Create section

For this section, a single change is proposed: replacing the story creation button with its equivalent AAC symbols, in order to foster greater appropriation of the system and ensure consistency in symbol use throughout the interface (see Figure 43).

It is also proposed to display the selected pictograms at a larger size during the story generation phase, while they accompany the waiting animation.

Figure 43. Adjustment in the create section: Substitute the text button for ARASAAC pictograms.

Story section

For this section, the addition of an opening page is proposed, presenting the story title in pictograms and highlighting the user's name as co-creator of the story (see Figure 44). A reorganization of the page layout is also suggested, giving pictograms greater visual prominence and introducing spacing between them that corresponds to the text they represent. When a word has no pictogram equivalent, a blank space is left beneath it. The illustration is proposed to expand and occupy more of the page, while the navigation controls are repositioned to the bottom of the tablet for easier access (see Figure 45).

Finally, a closing page is proposed, featuring the appropriate pictogram and a question inviting the AAC user to indicate whether they enjoyed the story (see Figure 46).

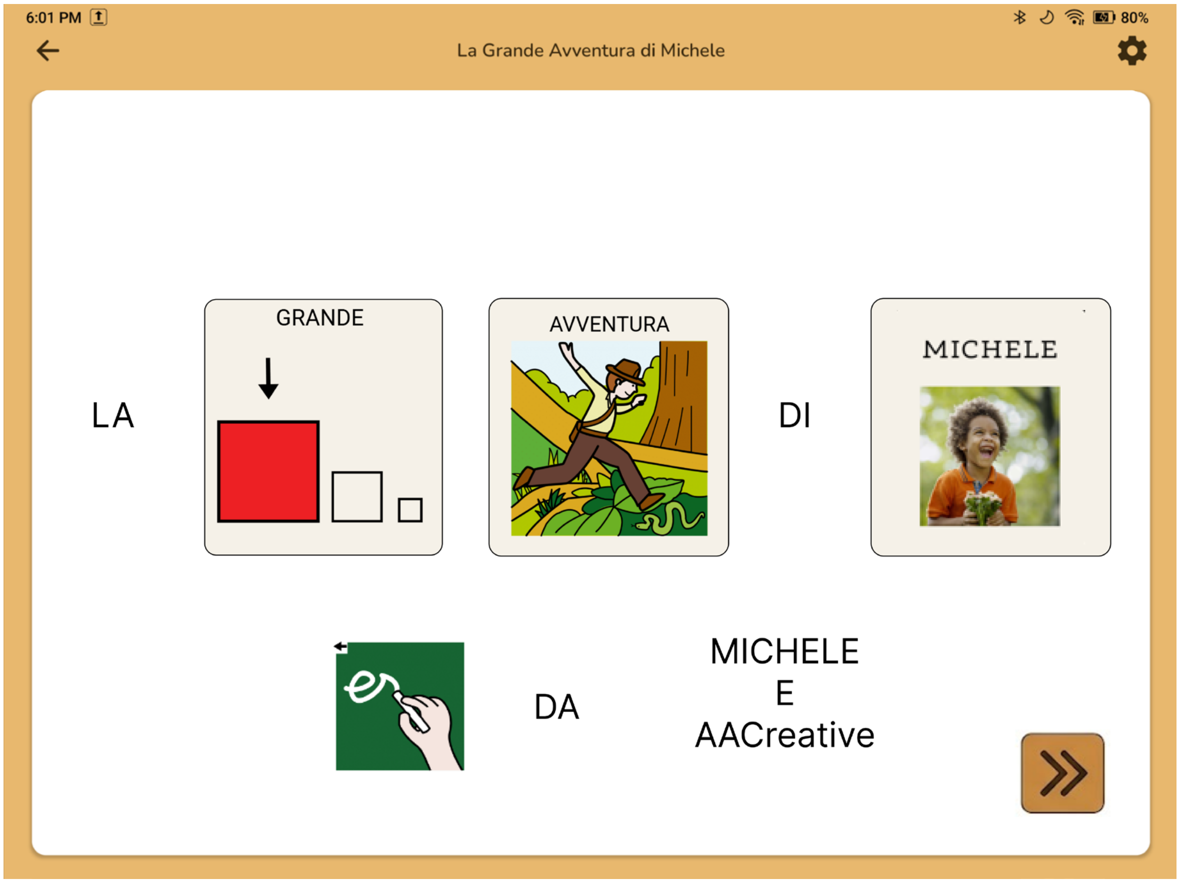


Figure 44. First page of the story presents the title and the authors.

Figure 45. Proposal for rearrangement of the elements in the storybook (made with Gemini AI)

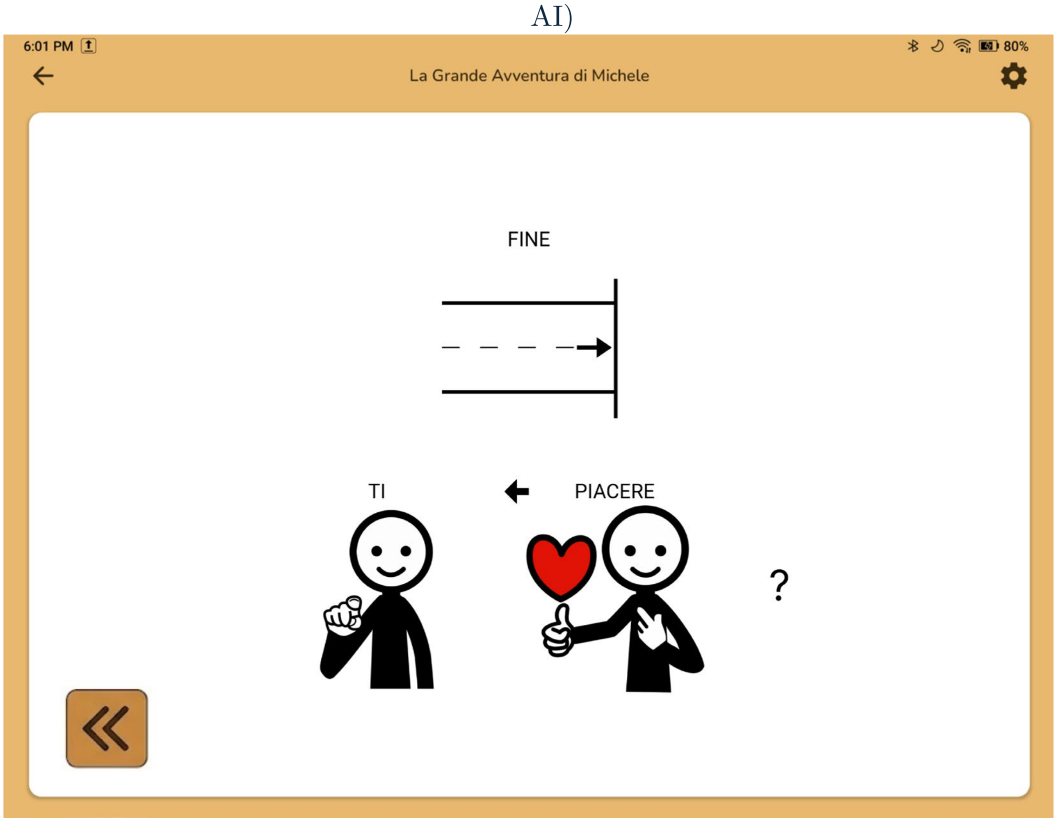


Figure 46. Page to close the story.

Prompt adjustments for story generation

Based on repeated interaction with the application, the following adjustments are proposed for the story generation prompt:

- ✓ Maintain visual consistency across illustrations between pages. It was observed that character features such as hair color, hairstyle, and clothing sometimes change without apparent reason, and accessories such as glasses appear or disappear from one page to the next. In some cases, key characters undergo significant visual changes that affect narrative coherence and comprehension. The same issue was noted with setting elements: for example, a castle depicted on one page appeared differently on the following page.
- ✓ Use the descriptions of custom cards as relevant input for story generation.
- ✓ Apply age-appropriate sentence structures and vocabulary calibrated to the AAC user's developmental level. The Flesch-Kincaid Grade Level scale may serve as a useful reference for this purpose (see Personalized storybooks using generative AI: Autihero section).
- ✓ Where the story length allows, provide a more coherent closure — for example, by resolving actions that were introduced earlier rather than introducing new situations on the final page.
- ✓ Maintain thematic coherence across pages. If a story is set around a birthday party, all pages should remain connected to that event unless a narrative reason for shifting focus is provided.
- ✓ Ensure that all cards selected as story input are reflected in the generated narrative.
- ✓ Introduce greater variety in story titles, as many of the generated titles contain the word "adventure."
- ✓ When a proper name appears in the story without a corresponding custom card, avoid generating a pictogram equivalent, or substitute it with a generic card such as "the girl" or "the boy."

- ✓ When emotions are present in the story, ensure that their inclusion is narratively justified — that is, any change in a character's emotional state should correspond to an event described within the story.

6.2. ADJUSTMENTS TO THE PHYSICAL SETUP

To establish a stronger connection with current therapeutic practice, a reorganization of the card frames is proposed, grouping them into categories of words or phrases that respond to the basic questions needed to select the elements of a story or sentence: Who? What? Where? and When? This connects partially with the Colorful Semantics method[28]; however, a color coding system closer to the Fitzgerald Key[28] is preferred in order to maintain consistency with other AAC systems that users may already be familiar with. A vertical orientation of these categories is also proposed, once again to align with the most widely used AAC applications, which typically organize their content from left to right. The intention is for these categories to be arranged at the therapist's discretion, allowing them to reorder the columns and decide how many to include. For example, a How? category could be added to accommodate adverbs, adjectives, and emotional states, enabling more detailed descriptions of situations and enriching the resulting narrative. It is worth noting that the What? category is primarily oriented toward verbs, though it is intentionally left open to also accommodating events or situations in a broader sense.

The idea is that the therapist or caregiver can define as many columns as needed, and that the AAC user can select one or more elements from each category without being required to fill all of them, depending on the desired level of narrative complexity. For example, a story may involve more than one character and more than one action but is less likely to require multiple locations or temporal elements. It is important that the resulting story reflects the **order in which elements are selected**, as this sequencing may give rise to meaningfully different narratives depending on the category involved.

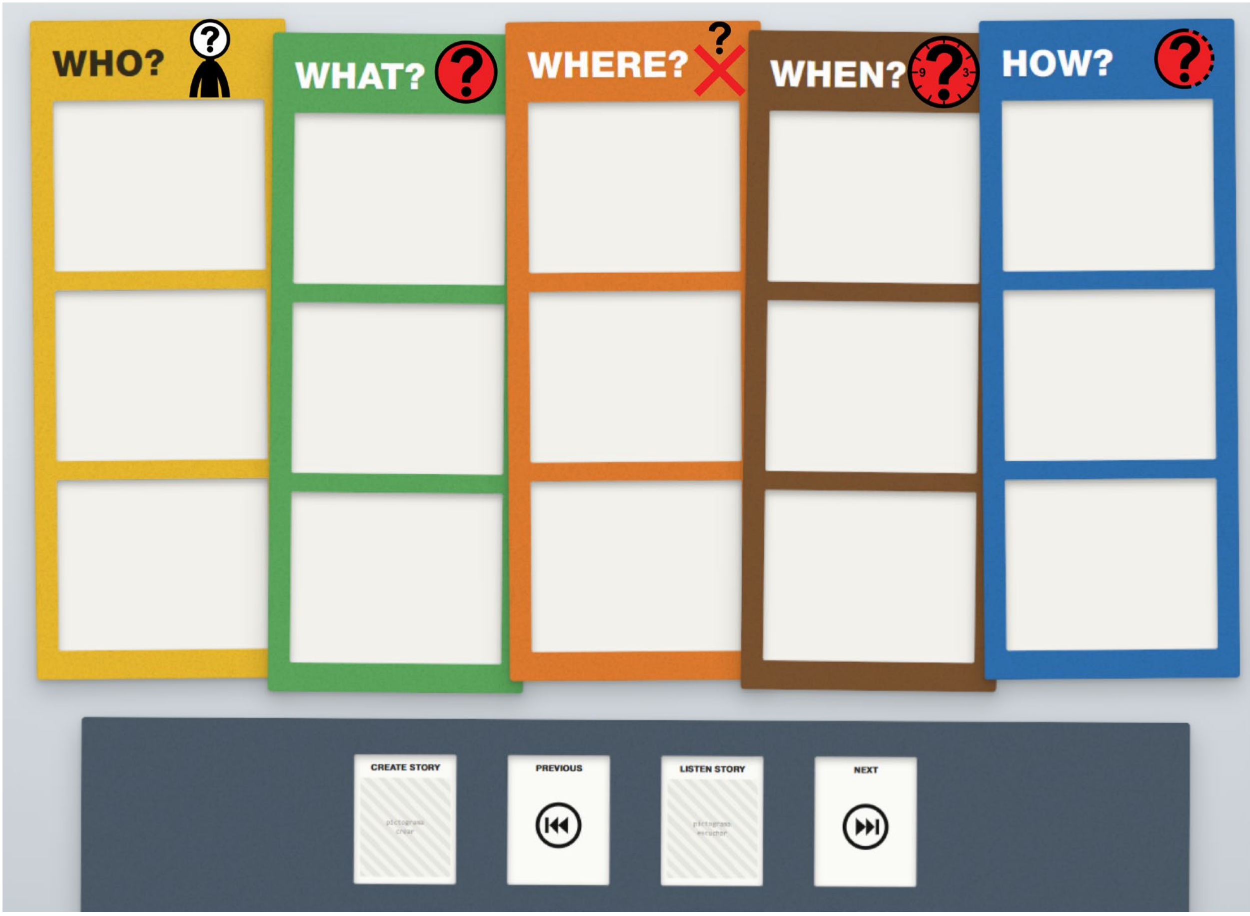


Figure 47. Proposal for the frames that would contain the pictograms.

Regarding to the scanner, the guidance received during the co-creation session indicates that a horizontal hand orientation is easier to achieve for users with a degree of motor impairment that makes touchscreen interaction difficult, but who still retain some ability to grip the reader. For this case, some of the components used so far can be adjusted: a smaller ESP32 board and a smaller battery can be organized within a housing that accommodates the components in a way that makes card scanning intuitive while keeping the vibration motor close to the palm of the hand. Some visual references are shown in Figure 48. However is important to highlight that this approach will depend on a deeper knowledge on the particular AAC user.

Figure 48. Different ways of gripping that support motor impairment. Source: Pinterest

6.3. ADDITIONAL PROPOSALS

The intention is not to replace current therapeutic practice but to enrich it by offering an additional interaction tool. For example, the AAC user could first select the cards that will generate the story and then be asked to narrate their own version of it using whatever communication means they have available. The story would then be generated through the system using the same cards, allowing both versions to be compared. Over time, this exercise could gradually expand the user's narrative resources, as they would have the opportunity to observe stories that, starting from the same elements, produce potentially different outcomes.

An alternative in the same direction would be to create dice featuring AAC symbols (see Figure 49) organized around the same narrative question categories defined earlier. Each dice could correspond to a category; for example, a characters dice, a verbs dice, a places dice, and so on. Each face of the dice would display the corresponding card with its NFC tag and could be scanned in the same way as the existing cards. As observed in the recorded therapy sessions, the act of rolling dice introduces a more playful quality to the experience.

Figure 49. AAC dices to create random stories first by the AAC user, then by the system. (image generate by Gemini AI)

7. LIMITATIONS AND FUTURE WORK

Due to time constraints, it has not been possible to test the device directly with AAC users with the characteristics described in this document. While expert guidance has been available throughout the process, a direct session with the primary user, both for the application and for the physical interaction layer, would undoubtedly generate new insights that could significantly expand the design possibilities. Formal user testing with minimally verbal children remains a necessary next step before any conclusions can be drawn about the effectiveness of the system in real therapeutic contexts.

The proposal does not extend to users with severe motor impairments. As noted in previous sections, those cases require a high degree of individualization and adaptation that goes beyond the scope of this work. In scenarios involving significant motor limitations, dedicated access devices such as eye trackers or large switches enabling indirect selection through screen scanning would likely be more appropriate and effective than the tangible card-based interaction proposed here.

The interaction, while brief, requires the AAC user to sustain continuous attention across three sequential phases: selecting the input cards, waiting for the story to be generated, and reading or listening to the complete story. This attention demand may present a barrier for some users, particularly those with significant attentional difficulties or low tolerance for delayed feedback. Similarly, understanding the causal relationship between physical actions and their effects on the screen requires a learning curve that has not yet been formally evaluated.

The physical prototype in its current form remains a functional sketch rather than a finished product. The housing of the components, the durability of the NFC tags, and the ergonomics of the controller have not been evaluated under conditions of prolonged or repeated use, an evaluation that would need to involve additional professional profiles such as occupational therapists.

From a narrative generation perspective, the AI story output still needs to find a better balance between simplicity and coherence across pages. The consistency of illustrations, the logical structure of the narrative, and the appropriate calibration of linguistic complexity to the user's developmental level are areas that require further prompt engineering and, ideally, clinical validation. The prompt adjustments proposed in this thesis are design hypotheses rather than validated solutions.

Finally, the system has been designed and evaluated primarily within an Italian clinical context, in collaboration with the Istituto Neurologico Carlo Besta. Its transferability to other cultural, linguistic, and institutional contexts has not been assessed and represents an open question.

Future Work

Several directions for future development emerge from the limitations and findings of this thesis. On the hardware side, using the ESP32-S3 as a more compact alternative to the ESP32-WROOM-32 currently used could significantly reduce the size of the controller and make it more ergonomic for users with different motor profiles. The form and material of the controller housing also remain to be explored with real users, particularly those with spasticity or reduced range of movement, for whom grip comfort and scanning posture are critical design variables.

A relevant direction for future work would be to systematically evaluate the impact of the prompt adjustments proposed in this thesis on narrative coherence, linguistic appropriateness, and illustration consistency, and to assess how therapists and children respond to the resulting stories across different developmental levels and therapeutic objectives. Related to this, a configuration layer that allows therapists to adjust generation parameters directly within the application is something that the student of computer science involved in this project is working on.

The story dice concept introduced in the redesign section also merits further exploration as a standalone development direction. Dice featuring AAC symbols organized by narrative category, each face equipped with an NFC tag that can be scanned in the same way as the existing cards, would introduce the element of chance and playfulness observed in recorded therapy sessions, while maintaining continuity with the physical interaction model developed in this thesis.

Caregivers represent another underexplored user group. While this thesis focused primarily on the therapeutic context, caregivers may find value in the system in quite different ways, for example, using it at home to prepare a child for upcoming events through anticipatory narratives, to revisit shared experiences, or simply as a creative activity that does not require clinical mediation. How caregivers appropriate the tool outside the therapy room, and what adjustments would make it more accessible and useful in domestic contexts, could open design directions that are currently outside the scope of this work.

Finally, exploring the system's use in school settings, home environments, and other countries with different AAC traditions and therapeutic cultures would significantly broaden the understanding of its potential and its limits. Alongside this, developing a specific evaluation protocol for AAC storytelling tools, one that measures not only usability but also impact on narrative competence, engagement over time, and therapist perception of clinical utility, would provide a more robust framework for assessing the contribution of systems like AACreative to the field.

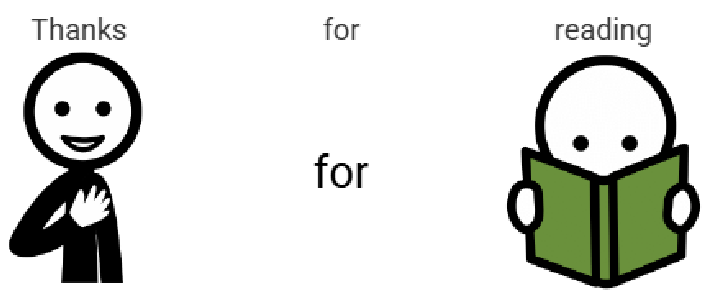
8. CONCLUSIONS

This thesis presented the design of a tangible interaction layer for AACreative, an existing AI-powered storytelling application for minimally verbal children who use AAC systems. The motivation for introducing a physical interaction layer responded to two converging considerations: on one hand, the body of research on Tangible User Interfaces, which demonstrated that physical manipulation of objects supports cognitive engagement, promotes a sense of authorship, and offers forms of embodied cognition that graphical interfaces alone cannot provide; on the other hand, the recognition that many minimally verbal children can also present some degree of motor impairment, which can make touchscreen interaction unreliable, fatiguing, or altogether inaccessible. Through a user-centered and iterative process, combining exploratory digital research, semi-structured interviews with speech-language pathologists and caregivers, observation of therapy sessions, and a co-creation session with clinical experts, the work sought to understand whether and how physical interaction with NFC-tagged cards could extend the expressive and therapeutic potential of a system originally designed for touchscreen use.

The user research phase produced a coherent and grounded picture of the therapeutic context. Across all methods, experts consistently emphasized that personalization is not an enhancement but a prerequisite for effective therapeutic material with this population: stories and resources that incorporate the child's name, familiar people, and everyday contexts generate a significantly higher degree of engagement and identification than generic content. Producing this level of personalized material has traditionally demanded considerable time and effort from therapists and caregivers. In this regard, generative AI represents a meaningful opportunity. By allowing caregivers to input contextual details, systems like AACreative can generate personalized narrative and visual content at a scale and speed that would otherwise be unfeasible, lowering the production burden without sacrificing the contextual relevance that makes these materials therapeutically effective. The research also surfaced the persistent gap between digital AAC tools and the physical materials that continue to anchor many therapeutic practices. For example, printed cards arranged across different spaces carry affordances that screens do not easily replicate, including spatial arrangement, tactile feedback, and presence in the environment as persistent visual anchors. This gap informed the central design proposition of the thesis.

Two sketches in hardware were developed and brought into dialogue with clinical practice, generating insights that informed a concrete set of proposed adjustments to both the physical setup and the application interface. Between the proposals are the reorganization of card frames into narrative question categories, color coding aligned with grammatical roles, an integrated flow for custom card creation and tag association, haptic feedback at critical interaction moments, and a series of prompt refinements aimed at producing more coherent and developmentally appropriate stories. The resulting prototype was not a finished product but a design proposition, a tangible layer that positioned the AAC user as an active author of AI-generated narratives, mediated by physical objects rather than a screen, and embedded within the relational dynamics of therapeutic and caregiving contexts.

While questions remain about card manipulation, user adaptability, and AI-generated stories, this work ensures they can now be asked with much greater precision. Future testing with AAC users will have a much clearer starting point because of this process. Ultimately, this thesis delivered something prior and necessary: a practical design proposition grounded in real clinical contexts, built through dialogue with experts, and made concrete through hardware sketches. There is still a lot of territory to explore with Tangible User Interfaces in AAC, but this project successfully maps out one specific, viable combination.



BIBLIOGRAPHY

[1] C. Basil, “What is AAC? - ARASAAC.” Accessed: Jun. 02, 2026. [Online]. Available: https://arasaac.org/aac/en

[2] Y. Elsahar, S. Hu, K. Bouazza-Marouf, D. Kerr, and A. Mansor, “Augmentative and Alternative Communication (AAC) Advances: A Review of Configurations for Individuals with a Speech Disability,” *Sensors*, vol. 19, no. 8, p. 1911, Apr. 2019, doi: 10.3390/s19081911.

[3] “AAC symbols - ARASAAC.” Accessed: Jun. 22, 2026. [Online]. Available: https://arasaac.org/pictograms/search

[4] H. Tager-Flusberg and C. Kasari, “Minimally Verbal School-Aged Children with Autism Spectrum Disorder: The Neglected End of the Spectrum,” *Autism Res.*, vol. 6, no. 6, pp. 468–478, Dec. 2013, doi: 10.1002/aur.1329.

[5] A. Kaiser, “Comprehensive Communication Intervention for Minimally Verbal Children With Autism,” clinicaltrials.gov, Clinical trial registration NCT02291172, Jul. 2018. Accessed: Jun. 12, 2026. [Online]. Available: https://clinicaltrials.gov/study/NCT02291172

[6] “Assistive products for persons with disability — Classification and terminology,” International Organization for Standardization (ISO), Geneva, International Standard ISO 9999:2022, 2022. [Online]. Available: https://www.iso.org/standard/60547.html

[7] “Priority Assistive Products List (APL),” World Health Organization (WHO), Geneva, Technical Report, 2016. [Online]. Available: https://iris.who.int/server/api/core/bitstreams/665850a1-6207-4c68-b1d2-b3b507ee77cd/content

[8] T. Griffiths and A. Addison, “Access to communication technology for children with cerebral palsy,” *Paediatr. Child Health*, vol. 27, no. 10, pp. 470–475, Oct. 2017, doi: 10.1016/j.paed.2017.06.005.

[9] H. Koester, S. Koch Fager, J. Gormley, E. Jakobs, K. Johnsen, and J. Brumberg, “Supporting effective alternative access for individuals with physical disabilities: state of the science, emerging technologies, and future research directions,” *Augment. Altern. Commun.*, vol. 41, no. 3, pp. 304–317, Jul. 2025, doi: 10.1080/07434618.2025.2499676.

[10] G. Soto and E. Hartmann, “Analysis of narratives produced by four children who use augmentative and alternative communication,” *J. Commun. Disord.*, vol. 39, no. 6, pp. 456–480, Nov. 2006, doi: 10.1016/j.jcomdis.2006.04.005.

[11] G. Soto, B. Yu, and S. Henneberry, “Supporting the development of narrative skills of an eight-year old child who uses an augmentative and alternative communication

device," *Child Lang. Teach. Ther.*, vol. 23, no. 1, pp. 27–45, Feb. 2007, doi: 10.1177/0265659007072145.

[12] N. M. Almubark, T. D. Spencer, and M. E. Foster, "AAC narrative intervention for children with autism," *Augment. Altern. Commun.*, pp. 1–14, Jan. 2025, doi: 10.1080/07434618.2024.2433959.

[13] J. Wu and Y. Xie, "Augmentative and Alternative Communication in the Context of Dramatic Storytelling: A Single-Case Study of a Pre-schooler With a Cochlear Implant," *Front. Educ.*, vol. 7, p. 806665, Apr. 2022, doi: 10.3389/feduc.2022.806665.

[14] N. Grove, *Using Storytelling to Support Children and Adults with Special Needs: Transforming Lives Through Telling Tales*. Routledge, 2013.

[15] H. Ishii and B. Ullmer, "Tangible bits: towards seamless interfaces between people, bits and atoms," in *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, Atlanta Georgia USA: ACM, Mar. 1997, pp. 234–241. doi: 10.1145/258549.258715.

[16] B. Ullmer and H. Ishii, "Emerging frameworks for tangible user interfaces," *IBM Syst. J.*, vol. 39, no. 3.4, pp. 915–931, 2000, doi: 10.1147/sj.393.0915.

[17] B. Hengeveld, C. Hummels, K. Overbeeke, R. Voort, H. Van Balkom, and J. De Moor, "Tangibles for toddlers learning language," in *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, Cambridge United Kingdom: ACM, Feb. 2009, pp. 161–168. doi: 10.1145/1517664.1517702.

[18] M. Resnick *et al.*, "Digital manipulatives: new toys to think with," in *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '98*, Los Angeles, California, United States: ACM Press, 1998, pp. 281–287. doi: 10.1145/274644.274684.

[19] J. Lee, K. Lee, I. Hwang, S. Park, and Y.-H. Kim, "AutiHero: Engaging Parents in Creating Personalized, Multi-path Social Narratives for Autistic Children," 2025, *arXiv*. doi: 10.48550/ARXIV.2509.17608.

[20] "Flesch Reading Ease and the Flesch Kincaid Grade Level," Readable. Accessed: Jun. 25, 2026. [Online]. Available: https://readable.com/readability/flesch-reading-ease-flesch-kincaid-grade-level/

[21] "International language standards | Cambridge English." Accessed: Jun. 25, 2026. [Online]. Available: https://www.cambridgeenglish.org/exams-and-tests/cefr/

[22] X. Hu *et al.*, "NarrAid: Supporting Storytelling of People with Aphasia via Generative Visual Scene Displays," in *Proceedings of the Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems*, in CHI EA '26. New York, NY, USA: Association for Computing Machinery, Apr. 2026, pp. 1–11. doi: 10.1145/3772363.3798625.

[23] E. A. Beccaluva *et al.*, “Grasping the Grammar with Moovy: a Tangible User Interface to Train Linguistic Skills in Children,” in *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, in CHI EA ’22. New York, NY, USA: Association for Computing Machinery, Apr. 2022, pp. 1–7. doi: 10.1145/3491101.3519733.

[24] L. Shapiro and S. Spaulding, *The Routledge Handbook of Embodied Cognition*. Taylor & Francis, 2024.

[25] B. J. Hengeveld, “Designing LinguaBytes : a tangible language learning system for non- or hardly speaking toddlers,” Phd Thesis 1 (Research TU/e / Graduation TU/e), Technische Universiteit Eindhoven, Eindhoven, 2011. doi: 10.6100/IR715535.

[26] “Avaz Full-featured AAC app - Features,” Avaz Inc. Accessed: Jun. 25, 2026. [Online]. Available: https://avazapp.com/features/

[27] Avaz, Inc., *Avaz Step Up Demo*, (Mar. 31, 2025). Accessed: Jun. 25, 2026. [Online Video]. Available: https://www.youtube.com/watch?v=r70xIYPbMm8

[28] Jayr Alencar Pereira, “A Method for Adapting Large Language Models for Communication Card Prediction in Augmentative and Alternative Communication Systems,” UNIVERSIDADE FEDERAL DE PERNAMBUCO, 2023. doi: 10.13140/RG.2.2.20576.46088.

[29] AssistiveWare, “Proloquo2Go User Manual.” 2015. Accessed: Jun. 25, 2026. [Online]. Available: https://orin.com/access/docs/Proloquo2Go-4-Manual.pdf

[30] “Proloquo2Go - AAC app with symbols - AssistiveWare.” Accessed: Jun. 25, 2026. [Online]. Available: https://www.assistiveware.com/products/proloquo2go

[31] “AAC Symbols and shared resources - ARASAAC.” Accessed: Jun. 22, 2026. [Online]. Available: https://arasaac.org/

[32] “PECS®: An Evidence-Based Practice -,” Pyramid Educational Consultants USA. Accessed: Jun. 22, 2026. [Online]. Available: https://pecsusa.com/pecs/

[33] “Tutorial CAA – ‘How to recognize the color keys of pictograms’, prepared by ARASAAC - Aula abierta de ARASAAC.” Accessed: Jun. 29, 2026. [Online]. Available: https://aulaabierta.arasaac.org/en/tutorial-caa-how-to-recognize-the-keys-of-color-of-the-pictograms