

The Acting Argument Paradigm: An Innovative Approach to Speech Therapy Using the Moovy TUI

ELEONORA AIDA BECCALUVA, Department of Psychology, Università degli Studi di Milano-Bicocca, Milan, Italy and Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy
MATHYAS GIUDICI, Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy

ELEONORA PASQUA, CRC Centro Ricerca, Rome, Italy

FABRIZIO AROSIO, Department of Psychology, Università degli Studi di Milano-Bicocca, Milan, Italy

FRANCA GARZOTTO, Department of Psychology, Università degli Studi di Milano-Bicocca, Milan, Italy and Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy

Developmental Language Disorder (DLD) is characterized by persistent and significant language impairments. Children with DLD may experience difficulties in grammar, phonology, vocabulary, and specific morphosyntactic structures, such as clitic pronouns and passive constructions. To address these challenges, we propose an innovative treatment approach termed the Acting Argument Paradigm (AAP). Grounded in psycholinguistic theories, the core idea of this approach is to translate the abstract linguistic concept of argument movement into physical actions that actively engage children during language therapy. To operationalize the AAP, we developed a Tangible User Interface, namely Moovy, that enables children to manipulate physical objects and fosters a multisensory learning experience. This article presents the AAP and Moovy and discusses the results of a 6-week experimental study involving children ($N=30$) with DLD. Preliminary findings indicate Moovy's potential as an effective tool for speech therapy programs, offering promise for improving the linguistic skills of children with DLD.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**; **Interaction paradigms**; *Accessibility technologies*; • **Applied computing** → **Interactive learning environments**;

Additional Key Words and Phrases: Tangible User Interfaces (TUI), children, smart toy, speech-therapy, Developmental Language Disorder (DLD), psycholinguistics

Eleonora Aida Beccaluva and Mathyas Giudici contributed equally to the article.

Authors' Contact Information: Eleonora Aida Beccaluva, Department of Psychology, Università degli Studi di Milano-Bicocca, Milan, Italy and Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy; e-mail: e.beccaluva@campus.unimib.it; Mathyas Giudici (corresponding author), Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy; e-mail: mathyas.giudici@polimi.it; Eleonora Pasqua, CRC Centro Ricerca, Rome, Italy; e-mail: e.pasqua@crc-balbuzie.it; Fabrizio Arosio, Department of Psychology, Università degli Studi di Milano-Bicocca, Milan, Italy; e-mail: fabrizio.arosio@unimib.it; Franca Garzotto, Department of Psychology, Università degli Studi di Milano-Bicocca, Milan, Italy and Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy; e-mail: franca.garzotto@polimi.it.



This work is licensed under Creative Commons Attribution-NonCommercial-NoDerivatives International 4.0.

© 2026 Copyright held by the owner/author(s).

ACM 1557-7325/2026/8-ART45

<https://doi.org/10.1145/3820893>

ACM Reference format:

Eleonora Aida Beccaluva, Mathyas Giudici, Eleonora Pasqua, Fabrizio Arosio, and Franca Garzotto. 2026. The Acting Argument Paradigm: An Innovative Approach to Speech Therapy Using the Moovy TUI. *ACM Trans. Comput.-Hum. Interact.* 33, 4, Article 45 (August 2026), 29 pages.
<https://doi.org/10.1145/3820893>

1 Introduction

Children with **Developmental Language Disorder (DLD)** have severe and persistent problems with language comprehension and production despite age-appropriate nonverbal cognitive abilities [47]. Although DLD is a heterogeneous disorder, some areas of the linguistic competence of children with DLD are particularly impaired. These generally include phonology [47], vocabulary [23], syntax, and morphosyntax [70]. Morphosyntactic deficits typically manifest as difficulties in producing and comprehending passive sentences (e.g., “*La mela è mangiata da Luca.*” [“*The apple is eaten by Luca.*”]) and object relative clauses (e.g., “*La mela che Luca sta mangiando è rossa.*” [“*The apple that Luca is eating is red.*”]). In Italian, morphosyntactic deficits also affect the production of **direct Object (DO)** clitic pronouns, whose omission is considered a reliable clinical marker of DLD. DO clitic pronouns are short morphemes that replace a DO and attach closely to a verb or appear immediately next to it—an example is “*Luca compra una mela e poi **la**^(clitic pronoun) mangia.*” [“*Luca buys an apple and then he **it**^(clitic pronoun) eats.*”].

While Italian-speaking children with typical language development acquire passive sentences and object relative clauses around the age of five, and begin using clitic pronouns at around 2 years of age, children with DLD often omit clitics and continue to experience difficulties with passives and relative clauses even at school age [4, 7, 8, 14, 36].

In Italian, sentences with clitics, passive sentences, and object relative clauses share a challenging form of *argument movement*, namely the movement of an internal verbal argument from its canonical post-verbal position within the verb phrase (e.g., “*Luca mangia **la mela**.*” [“*Luca eats the **apple**.*”]) to the left periphery of the sentence (e.g., in a passive form, “***La mela** è mangiata da Luca.*” [“***The apple** is eaten by Luca.*”]), and in the clitic form, “*Luca **la**^(clitic pronoun) mangia.*” [“*He **it**^(clitic pronoun) eats.*”]) [5, 17].

The central idea underpinning our research was to propose a new therapeutic approach that *translates the grammatical concept of argument movement into a physical action* unfolding before children’s eyes and actively engaging them. This idea evolved into a novel *intervention paradigm*, the **Acting Argument Paradigm (AAP)**, and into *Moovy*, a **Tangible User Interface (TUI)** for linguistic intervention. *Moovy* operationalizes the AAP by allowing children to reify argument movement through the manipulation and movement of physical objects.

The development of the AAP and *Moovy* is grounded in the extensive literature on psycholinguistics as well as in theories of Embodied Cognition [64, 80] and Sensory Integration [66], which posit a link between cognitive processes, movement, and multisensory stimulation and provide a theoretical foundation for TUIs. Our approach is also underpinned by theories and empirical studies on how the brain processes language and action [57]. They indicate that handling physical objects recruits the same basal ganglia circuits underlying syntactic computation [69], a phenomenon that is not observed during generic hand motion [20]. These results provide a principled neural basis for *embodied* language intervention that involves intentional, goal-directed object manipulation and movement through TUIs.

In addition to introducing the AAP and describing the design and implementation of *Moovy*, this article presents the results of an experimental study involving 30 children with DLD. The participants took part in a 6-week morphosyntactic training program using *Moovy*. We assessed both the

effectiveness of Moovy and the validity of the underlying AAP paradigm. Although preliminary, our findings provide promising evidence of the potential of AAP and Moovy as effective tools for speech therapy, and underscore the value of integrating body-mind and psycholinguistic theories in the design of new TUI-based solutions for individuals with language impairments.

The remainder of the article is organized as follows. Section 2 presents the theoretical underpinnings of our research and reviews the state of the art on relevant interactive technologies for individuals with language impairments, with particular emphasis on TUIs. Section 3 describes the AAP. Section 4 presents Moovy, detailing its features and functionalities. Section 5 presents the methodology of our empirical study. Section 6 reports the main results of the study, while Section 7 discusses their implication. Section 8 synthesizes our findings and offers concluding remarks. Finally, Appendix A offers a comprehensive description of the linguistic theory on the argument structure of sentences, clitic pronouns, and passive sentences.

2 Related Work

This section first introduces key linguistic concepts related to argument movement, and then presents the theoretical background underlying the AAP and Moovy, focusing on the relationship between language and motor processes and on relevant theories underpinning TUIs for language learning. Subsequently, we review selected literature on the use of interactive technologies—particularly TUIs—in the domain of language intervention.

2.1 Arguments, Clitics Pronouns, and Passive Sentences

Within sentences, individual words are organized into hierarchically structured sequences, forming units known as *phrases*. *Arguments* are phrases that combine with a verb to encode participants in the event or state described by that verb [1]. Arguments may occupy canonical or displaced positions depending on the language’s morphosyntactic operations [1, 28]. In both Italian and English, the canonical order is subject–verb–object [22]. Consider the sentence “*The nice guy sends the letter*.”; the verb “*send*” is transitive and requires two arguments: an external one (associated with the subject position) and an internal one (associated with the object position). In the example, the “*nice guy*” is the external argument, while the “*letter*” is the internal argument.

However, verb arguments can also appear in uncanonical orders, depending on some grammatical requirements. Some constructions require *argument movement* to non-canonical positions. Among these are sentences with object clitic pronouns and passive sentences, which involve the displacement of the verb’s internal argument from its canonical post-verbal position to a pre-verbal position.

Object clitic pronouns are monosyllabic morphemes whose placement depends on verbal finiteness, namely, whether the verb form carries inflectional features (e.g., tense, person, and number) and can function as the main verb of an independent sentence. With finite verbs, clitic pronouns appear pre-verbally (proclitics), as in “*Il ragazzo carino **la**^(clitic pronoun) spedisce*.” [“*The nice guy **it**^(clitic pronoun) sends*.”]. In contrast, with non-finite verbs, they appear post-verbally (enclitics), as in “*La maestra dice al ragazzo di spedir**la**^(clitic pronoun)*.” [“*The teacher tells the boy to send **it**^(clitic pronoun)*.”] [6]. Since a clitic pronoun typically resumes a previously introduced discourse referent, the correct use of clitics is intrinsically challenging. It involves not only argument movement in the case of proclitics but also the integration of morphosyntactic agreement features, such as gender and number, with discourse-pragmatic constraints. Additionally, the complexity of mastering clitics may arise from the fact that these morphemes are phonologically weak and insubstantial, and therefore challenging to acquire [17]. An alternative hypothesis attributes the difficulties of mastering clitics to processing limitations, specifically insufficient cognitive resources [56]. This

interpretation is supported by the study by Grüter and Crago [35], which indicates that individuals with lower working-memory capacity produce fewer clitics. The multi-dimensional complexity of clitic pronouns makes them a robust clinical marker: Italian-speaking children with DLD often omit these forms or replace them with full noun phrases, well beyond the age at which typically developing peers have mastered them [17, 35].

Passive sentences are intricate syntactic structures characterized by an order of phrases that differs from the canonical one and by multistep argument movement resulting from a reorganization of grammatical functions. In a passive sentence, the internal argument of a transitive verb (object) becomes the subject of the sentence and moves to the pre-verbal position, the verb undergoes morphological transformations, and the original subject is demoted to an optional by-phrase. For instance, consider the sentence “*La madre nutre il bambino.*” [“*The mother feeds the child.*”]. In its passive counterpart “*Il bambino viene nutrito dalla madre.*” [“*The child is fed by the mother.*”], the object “*il bambino*” [“*the child*”] appears in the subject position; the verb undergoes morphological change, forming a structure composed by an auxiliary verb “*venire/essere*” [“*to be*”] followed by the past participle, both conjugated according to the number and gender of the newly established subject; and the original subject “*la madre*” [“*the mother*”] becomes the “*complemento d’agente*” [“*agent*”] [76]. The derivation of passive sentences can be described in terms of the smuggling hypothesis [13, 21]: the verb and the object first move together as a unit, thereby bypassing the external argument; only after this movement can the internal argument rise to the sentence’s left periphery, circumventing the usual blocking effect of the subject position.

Italian-speaking typically developing children show the ability to comprehend and produce passives between the ages of three and six, particularly under supportive conditions such as syntactic priming or with action verbs [48, 76]. However, performance varies depending on verb type, the presence of the by-phrase, and cognitive load. Since mastering the complexity of passive forms requires morphosyntactic competence, passive sentences pose particular challenges for children with DLD, who may continue to experience difficulties with passives beyond the age at which typically developing children generally show mastery of these structures [4, 7, 8, 14, 36].

Overall, both clitic pronouns and passive structures require children to coordinate phonological, morphosyntactic, and pragmatic information simultaneously. This is precisely why these linguistic constructs are central to the assessment of morphosyntactic abilities, are among the structures most systematically impaired in children with DLD [15], and represent a key target for linguistic intervention.

For a more comprehensive discussion of the above linguistic concepts, the reader is referred to Appendix A.

2.2 Language and Action

The relationship between language and action has long been a topic of extensive investigation [57]. Understanding how these domains interact offers key insights into brain evolution and cognitive architecture [52]. Research has traditionally focused on lexical-semantic aspects, such as how action-related words are grounded in sensory-motor systems [50, 60], while devoting less attention to syntactic structures. However, linguistic syntax and motor planning both seem to rely on overlapping neural substrates in Broca’s area, indicating a convergence between structural language processing and action organization [25, 26]. From an evolutionary standpoint, this overlap suggests that linguistic syntax may have emerged by exploiting pre-existing syntactic structures of the sensory-motor system [3, 27]. Both systems display hierarchical embedding and long-distance dependencies, with motor behaviors decomposing into goal-driven sub-events organized in structured sequences [40]. A number of theoretical proposals further reinforce this link. Gallistel [29] proposes that actions exhibit an internal argument structure analogous to that of verbs,

especially concerning subjects and objects. Moreover, the logical form of sentences has been interpreted as reflecting the sensorimotor processes involved in perceiving, planning, and executing real-world events [44].

The idea of a shared architecture underlying structural language and action processing is supported by neuroscientific evidence. Using **functional Magnetic Resonance Imaging (fMRI)**, Thibault et al. [69] compared the neural activation pattern in the same group of subjects during the performance of two different tasks: one involving syntactic processing and another involving the use of a physical tool. The fMRI results indicated that the basal ganglia, especially the striatum nucleus, are activated during both comprehending complex relative clauses and planning the actions to use the object correctly. This evidence suggests that linguistic processing and motor functions involve partially overlapping neural networks. Furthermore, upon examining the spatial distribution of neural activity in the two experimental conditions, the researchers observed that syntax tasks and tool use evoke similar patterns of brain activation. These findings support the hypothesis that these two types of tasks engage the same neural resources and rely on shared cognitive processes. When two functions share the same resources and activate similar cognitive processes, learning transfer typically occurs, which means that training in one function can positively impact another function that has not been directly trained.

Based on these findings, the researchers proposed the existence of a shared Supramodal Syntactic Function within the basal ganglia, activated during both syntactic processing and tool-use planning. This function is responsible for processing complex hierarchical structures, including embedded structures at both the motor and linguistic levels: for example, the incorporation of an instrument within a sequence of actions in the motor domain, and the embedding of one sentence within another in the linguistic domain. The concept of a Supramodal Syntactic Function provides a possible cognitive-neural basis for the hypothesis that language syntax and motor syntax are evolutionarily related.

Importantly, it seems that not all types of movements recruit shared neural mechanisms. The link between motor actions and linguistic syntax emerges specifically in contexts involving instrument manipulation, rather than generic hand motion. In Thibault et al. [69], syntax-related activations overlapped with neural activity elicited during tool-use tasks, but not during comparable tasks performed with the bare hand—highlighting the cognitive specificity of tool-mediated actions. This distinction is consistent with behavioral and kinematic studies indicating that dragging an object differs fundamentally from *grasping and manipulating* it, both biomechanically and neurally [20]. Only the latter engages sensory-motor processes that demand hierarchical structuring, intentional planning, and object-specific control [36]. This specificity also strengthens the evolutionary hypothesis that linguistic syntax emerged by leveraging and elaborating pre-existing motor syntax within the human sensory-motor system.

2.3 Theoretical Foundations of TUIs for Language Intervention

The potential of TUIs for language training and development is grounded in the same two theoretical perspectives that underpin TUI applications across many learning domains: Embodied Cognition [65, 80] and Sensory Integration [66].

Embodied Cognition emphasizes the importance of non-mental components and the intricate connection between the mind and body [65, 80]. It argues that cognitive processes arise from the body's interactions with the world, viewing learning as a holistic process involving both mental and physical aspects [80]. Sensory Integration [66] posits that individuals must integrate various environmental stimuli to form a complete mental picture of the world, suggesting that learning is facilitated through the integration of different sensory stimuli.

In addition, a particularly relevant theoretical framework for TUI-based language intervention comes from Glenberg [33] embodied account of comprehension, according to which language comprehension is enhanced by mapping linguistic symbols—including syntactic relations—onto action-based simulations grounded in perception and motor experience. This framework has been operationalized through the *Moved by Reading Intervention* [34]. This method comprises two stages. In the first stage, children manipulate physical toys to simulate the content of what they are reading, physically enacting the argument structure of sentences (i.e., who does what to whom). In the second stage, the children are taught to manipulate the toys in imagination. This approach has been shown to produce substantial gains in memory and comprehension. Although focused on reading rather than linguistic production, Glenberg's framework and findings provide further theoretical and empirical support for the use of embodied simulation in language intervention. They strengthen the motivation for adopting tangible interfaces in morphosyntactic training and support the rationale for our approach, in which a TUI is used to render abstract syntactic operations as concrete manipulations of physical objects.

2.4 Technologies and Language Impairments

Technology has become integral to the treatment of language impairments, offering a range of tools that support language development and rehabilitation. Among available technologies in this field, speech recognition tools stand out, aiding individuals with DLD by providing feedback on pronunciation, pitch, and speech. A study by Wang et al. [78] revealed that these tools were more effective in improving speech intelligibility compared to traditional speech therapy, emphasizing their potential to create personalized exercises tailored to individual needs. Assistive Communication Devices play a crucial role in facilitating communication for individuals with verbal limitations. These devices use pictures, symbols, and words to enhance communication outcomes, particularly for those with severe communication impairments, as indicated by an article by Beukelman and Mirenda [16].

The popularity of mobile applications and gaming technologies is on the rise as effective tools for treating linguistic disorders. These applications often leverage interactive and engaging games to help individuals practice language skills, including vocabulary and grammar [77]. Virtual Reality technology has been incorporated into language impairment treatment to simulate real-life scenarios, providing immersive and ecological experiences for language learning. Activities in virtual environments, such as ordering food at a restaurant or engaging in conversations with digital characters, offer opportunities for language training [45]. Additionally, wearable technology, such as smartwatches, serves as a valuable tool for language practice by offering reminders and prompts throughout the day. These features help individuals seamlessly incorporate language learning into their daily routines, therefore enhancing accessibility [55].

In developmental contexts, TUIs have been found useful for a wide range of learning goals, including those involving individuals with special needs. Several studies indicate that they can foster exploratory learning and support the transition from concrete to abstract understanding [62], enhance attention [10], improve problem-solving abilities [32], and promote sensory-motor skills [68]. Moreover, TUIs can promote collaboration skills, such as co-creation and collective problem-solving, by providing physical interfaces that allow multiple users to manipulate digital content together [63].

Although the application of TUIs to language-specific challenges has received comparatively less attention than other learning domains, the existing literature does provide examples that suggest their potential in addressing language-related issues. One notable TUI is the one designed by Sylla et al. [67] and named **Theory of Knowledge (TOK)**, which targets pre-literacy skills in five-year-old children. Using smart cards recognized by a platform, TOK engages children through open-ended

storytelling activities, fostering lexical knowledge and phonological awareness. Kiteracy [41] is a TUI addressing literacy development in children with Down Syndrome, employing both word-based and letter-based methods through a web-based application. Garzotto and Bordogna [30] and Garzotto and Gonella [31] proposed a TUI to promote cognitive and communication skills of non-verbal or minimally verbal children through low-cost, **Radio Frequency Identification (RFID)**-tagged paper-based objects, such as cards and posters, linked to multimedia content. An integrated authoring tool enabled caregivers to personalize these materials to address each child's needs. Several TUIs focus on specific aspects of language development. For instance, TangiLetter [81] combines TUI with speech-to-text technology to enhance English letter pair pronunciation in pre-school-aged children. RoyoBlocks [43] employs a plush companion and wooden blocks printed with common high-frequency words that young children are encouraged to recognize by sight, while providing a hands-on and engaging learning experience. PhonoBlocks [2] aids dyslexic children in learning English letter-sound pairs through color-coded cues. Beyond pre-literacy skills, tools like Kid Society [19] and LinguaByte [37] target receptive language development. Kid Society employs a board controller and physical tokens to improve vocabulary and concept classification. LinguaByte uses tangible multimedia storytelling tools to stimulate the communication skills of toddlers (with a developmental age between 1 and 4 years old) with multiple disabilities. Finally, expressive language development is addressed by tools like CreaTable [51], aiding individuals with aphasia in expressing themselves using interactive tables and printed words on cards.

3 AAP

Drawing on the theoretical background discussed in Section 2, we propose a new morphosyntactic intervention paradigm for Language Disorders, named the AAP. The AAP extends conventional intervention methods by integrating key principles from Embodied Cognition and psycholinguistics into language training. At its core, the AAP posits that effective language training can be achieved by engaging individuals in the manipulation of physical objects and in purposeful bodily movements that symbolize the rearrangement of syntactic elements. A central insight underlying the paradigm is the interconnectedness between bodily experience and language processing, as well as the parallelism between the organization of simple physical actions and the construction of sentence structures [60]. By aligning these cognitive processes, the AAP aims to bridge the gap between abstract linguistic concepts and embodied experience.

The AAP specifically targets linguistic constructs, such as clitic pronouns and passive sentences, that are characterized by complex hierarchical syntactic relations and require a mental rearrangement analogous to the motor movements performed by the body. Specifically, the AAP translates the *syntactic movement of arguments* involved in clitic pronouns and passive sentences into the manipulation and movement of physical objects. Additionally, by engaging brain areas associated with both linguistic and motor functions, the paradigm provides more focused and robust stimulation, potentially enhancing neural plasticity and promoting learning processes. Furthermore, the AAP envisions experiences that integrate diverse stimuli across multiple sensory channels and target a range of syntactic structures. This multifaceted, multisensory approach is intended to foster comprehension while supporting the cognitive processing of complex linguistic constructs.

3.1 Conceptual Scheme for AAP

This conceptual scheme in Figure 1 illustrates the flow of the AAP process, emphasizing the integration of linguistic concepts with embodied movements to foster a deeper and more embodied understanding of language. The conceptual scheme is built of:

(1) *Input: Linguistic Concept.*

The process begins with the conceptualization of a linguistic operation involved in the derivation of a complex syntactic structure. Central in this manuscript is the leftward movement of the verb's internal argument to a non-canonical position in passive and clitic constructions.

(2) *Motor Mapping: Association with Movements.*

The linguistic operation is associated with a specific body movement that represents the syntactic operation involved. For example, in the case of a passive sentence, a specific physical movement symbolizes the leftward movement of the verb's internal argument to a non-canonical position.

(3) *Embodied Exercise: Physical Engagement.*

Individuals perform the associated body movements while verbalizing or mentally processing the linguistic concept. This phase aims to synchronize cognitive processes with physical actions, reinforcing the connection between language and bodily experiences.

(4) *Feedback Loop: Iterative Learning.*

A feedback loop is established, allowing individuals to receive feedback on the accuracy of their movements and linguistic expressions. This iterative process supports continuous learning and refinement of the embodied mapping, enhancing the individual's grasp of the linguistic concept.

(5) *Application to Language Tasks.*

The embodied mapping learned through exercises is applied to broader language tasks, such as constructing and understanding different sentences requiring the same linguistic operation, or engaging in communicative scenarios.

(6) *Generalization: Transfer to Real-World Scenarios.*

Individuals generalize their embodied language skills to real-world scenarios, ensuring that the learned concepts are applicable beyond the training context. For instance, they may use embodied gestures during conversations, presentations, or language-related activities.

(7) *Adaptability: Individualized Approaches.*

The AAP is adaptable to individual needs and learning styles, allowing for personalized exercises and adjustments based on linguistic proficiency, motor skills, and cognitive abilities. For example, a user with limited motor abilities and early-stage linguistic impairment may be presented with simple visual stimuli and single-word association tasks, while a user with higher linguistic proficiency may engage in more complex exercises involving multimodal stimuli and sentence-level or narrative-based semantic processing. This flexibility enables the AAP to support personalized learning trajectories and to adapt over time as individual capacities evolve.

(8) *Outcome: Improved Language Competence.*

The ultimate goal of AAP is to enhance language competence by integrating the cognitive understanding of linguistic concepts with purposeful body movements. Improved language skills, comprehension, and communication are expected outcomes of regular engagement with the AAP.

4 Moovy

The unique aspect of Moovy lies in its approach to language learning through a TUI designed as a tabletop game. We advanced the preliminary presentation of Moovy [11], by turning the abstract concept of AAP into a tangible, motor action, game actively involving children in the learning process. This approach aligns with the principles of embodied learning, where physical actions are integrated with cognitive processes to enhance understanding and retention [65]. The game focuses

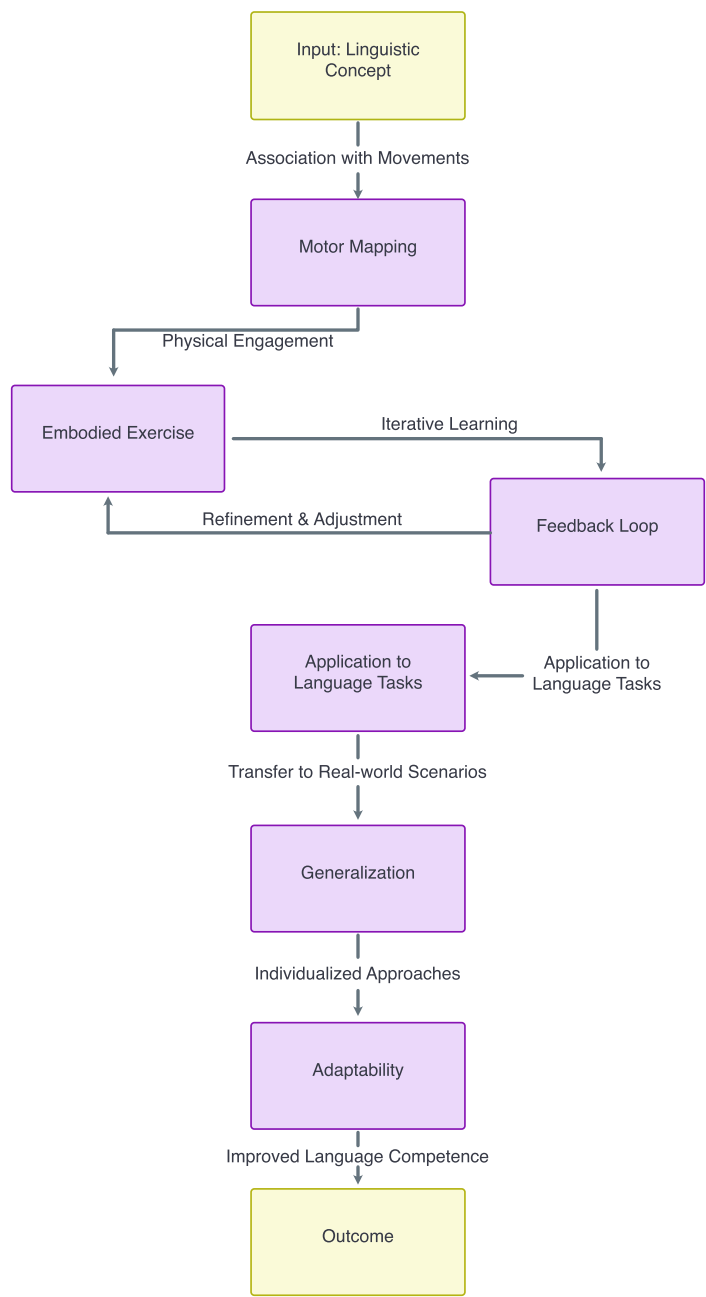


Fig. 1. The AAP process diagram.

on motor actions that metaphorically represent syntactic movement and provides a bridge between the abstract nature of language structures and the concrete, experiential world of the child. This hands-on, interactive approach not only aims to facilitate language learning but also to make the process enjoyable and engaging for the target age group (4–10 years old). The tangible interactive



Fig. 2. The Moovy elements on a table. On the left, the web app is on a tablet, and in the center, the table-top place, the Tangible Argument Objects, and the board.

design aligns with the principles of TUIs, providing a physical interface for manipulating digital information. Figure 2 shows Moovy and all its elements in a table.

4.1 Moovy in a Nutshell

The components of Moovy are designed to work together to create an engaging and effective learning experience. Here's a breakdown of the key components coded and mapped in Figure 3:

- *Table-top Panel (A)*: An A2-sized sheet of cardboard featuring representations of the 36 game boxes. This serves as the visual backdrop for the game, likely providing context or structure for the interactive elements.
- *Interactive Board (B)*: The main hub of Moovy, made of white plastic, housing essential components like an Arduino processor, RFID readers, a speaker, LED strips, cables, connectors, and a battery. The board is rectangular, resembling a grid with two columns and four rows. This grid likely corresponds to the layout of the game or the arrangement of interactive elements.
- *Tangible Argument Objects (C)*: Comprises 36 tangible objects representing arguments. Each object consists of a rigid black plastic base containing an RFID tag, with an argument representation attached on top in the form of a small toy, a picture, or a written word. The argument objects are versatile: thanks to the use of Velcro, different representations can be easily interchanged. Storage containers and a binder are provided for organizing and storing the objects.
- *Web App (D)*: A user-friendly web application that enables caregivers to create and manage children's profiles, initiate sessions, select specific activities, and monitor progress. This digital component enhances opportunities to customize each child's experience according to their specific needs.



Fig. 3. Moovy's table-top panel (A), interactive board (B), tangible argument objects (C), and web app (D).

The Interactive Board is highlighted as the primary interface, offering active feedback through LED illumination and auditory cues. The emphasis on minimalism ensures that, when not in active use, the interface does not distract children, maintaining focus on the learning activities. This comprehensive setup, combining physical and digital elements, aligns with the principles of TUIs' design [39, 53], providing a multimodal and interactive learning environment. The design requirements for the Moovy project are carefully crafted to ensure the effectiveness and appropriateness of the TUI in supporting morphosyntactic training for children with DLDs. The design principles followed are:

- *Activity Design*: The primary goal is to create activities that make participants feel at *ease, engaged, and motivated*. The creation process should consider the needs of children, allowing therapists to tailor activities to individual requirements. In addition, the activities should offer *multiple levels* of complexity, aligning with each child's performance level. This approach ensures that the TUI can be customized to suit varying degrees of difficulty, promoting a personalized learning experience. Finally, *modularity* is deemed essential, enabling the invention of a variety of activities, providing flexibility, and encouraging creativity in designing new tasks.
- *Movement*: The TUI is designed to *support movement* and action, aligning with the AAP. Physical materials are preferred, as they can be easily manipulated by children, and the integration of multiple senses through lights and sounds keeps children actively engaged.
- *Relevance*: Activities directly reflect relevant to *exercising complex morphosyntactic structures*, such as clitic and passive sentences. Clear feedback is provided when a child completes a task, ensuring a positive experience. Minimizing cognitive load during the activity is also considered.
- *Autonomy and Repetitiveness*: Allowing children to *act independently* without constant intervention promotes the development of a sense of *control over their actions*, supporting overall development. In addition, *repetition of activities* is emphasized for continuous progress in the target field. It not only facilitates improvement in the specific language domain but also enhances related areas, even those not directly targeted during training.
- *Simplicity and Predictability*: The TUI itself should be *as simple as possible* to avoid distractions from the task at hand. The design of the activity creation process must be predictable, clear, and intuitive, ensuring ease of use for therapists through a straightforward web app interface.

These requirements collectively contribute to the development of an effective, engaging, and customizable TUI that aligns with the specific needs and challenges of children with DLD. The emphasis on movement, relevance to language structures, autonomy, and repetition reflects a thoughtful and holistic approach to language training through interactive and embodied experiences.

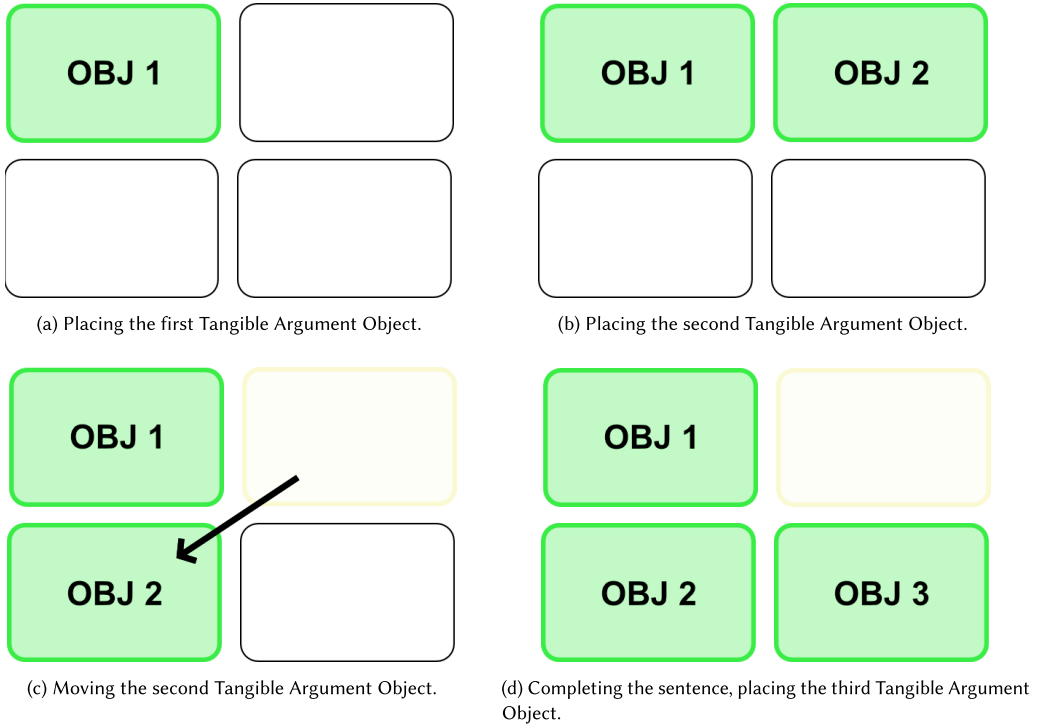
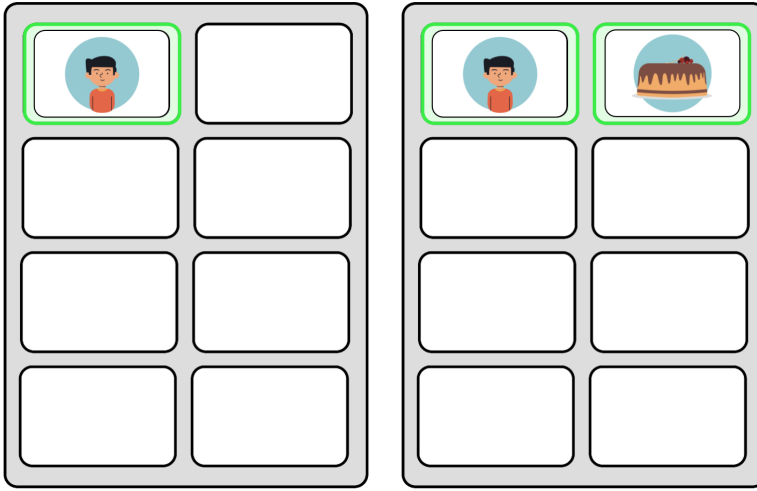


Fig. 4. Movement pattern of the Tangible Argument Objects (in the image labeled as “OBJ”) on the interactive board, with the *ghost trace* represented by the white highlighted area.

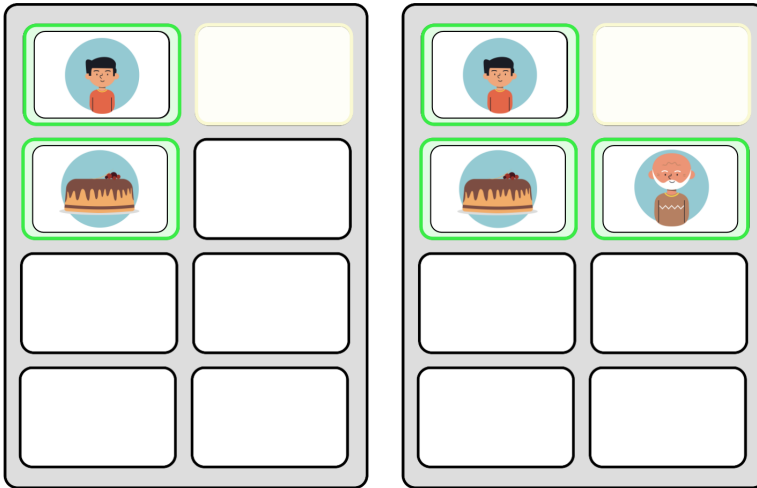
4.2 Playing the Moovy Game

Moovy activities focus on learning and practicing clitic pronouns and passive sentences. During gameplay, the interactive board produces a series of pairs of simple spoken sentences, each composed of a verb and its arguments and involving at least one of these morphosyntactic constructions. For each sentence, the child must select, among the Tangible Argument Objects placed on the Table-top Panel, those corresponding to the arguments of the sentence and position them in the appropriate locations on the board.

For example, upon hearing “*Prima il bambino mangia **la torta**.*” [“*First the child eats **the cake**.*”] as the first lead-in sentence, the child must select the “*child*” and “*cake*” Tangible Argument Objects from the tabletop panel and place them, respectively, in the upper-left and upper-right corners of the Interactive Board. Subsequently, the child hears a second sentence introducing the targeted morphosyntactic construction, for example: “*Poi **la torta** è mangiata dal nonno.*” [“*Then **the cake** is eaten by the grandfather.*”] (passive sentence). The child must move the “*cake*” Tangible Argument Object from the upper-right cell to the left cell in the row below, and then place the “*grandfather*” Tangible Argument Object in the adjacent cell on the right. A *ghost trace* visual effect is generated when the child enacts Argument Movement: the board marks the transition of a Tangible Argument Object from its canonical position to a different position by illuminating the vacated cell with a white light. Figure 4 depicts the main movement pattern described above, illustrating the child’s progressive placement of Tangible Argument Objects on the Interactive Board and the *ghost trace* effect, represented by the *white* highlighted area. Figure 5 shows this pattern through the previously presented example.



(a) Placing the first Tangible Argument Object (left) and second Tangible Argument Object (right) after the board played the first sentence “*Prima il bambino mangia **la torta**.*” [“First, the child eats **the cake**.”].



(b) Moving the second Argument Object (left) and placing a new Tangible Argument Object (right) after the board played the second sentence “*Poi **la torta** è mangiata dal nonno.*” [“Then **the cake** is eaten by the grandfather.”].

Fig. 5. Example of movement pattern, performed on the passive sentences: “*Prima il bambino mangia la torta. Poi la torta è mangiata dal nonno.*” [“First, the child eats the cake. Then the cake is eaten by the grandfather.”].

The interactive board produces audio sentences with a prosodic style intentionally designed to sound cheerful, engaging, and motivating for children, while still adhering to speech therapy guidelines for correct diction, rhythm, and fluency. At the same time, the sentences can be deceptively challenging, requiring the child to listen attentively, select the appropriate cards, and place them accurately on the board.

In addition to the ghost trace, the board provides several forms of immediate visual feedback. Green lights appear on the cells where the correct Tangible Argument Objects are placed, while red lights indicate errors. Finally, successful completion of a training session triggers a rainbow light effect over all the board's cells, accompanied by cheerful music, thereby reinforcing the session's positive and engaging conclusion of the Moovy experience.

Moovy's activities focus on two morphosyntactic constructs: clitic sentences and passive sentences. For each category, Moovy includes one "warm-up" set of sentence pairs (4 pairs) and 4 training sets (each containing 16 sentence pairs). Thus, each category comprises 68 pairs (136 sentences), resulting in a total of 272 sentences available for engagement. A linguistic training session typically comprises one set (i.e., 16 sentence pairs); using the Web App, caregivers can customize both the number of pairs and the specific pairs included in the session.

4.3 Technical Approach

Moovy employs a Client–Server architecture where therapists use a frontend interface to manage game activities, user information, and application settings, while the backend processes therapist actions and stores relevant data, including activities and sessions. The system features an interactive board that executes activities initiated by therapists and transmits session data to the backend through secure WebSocket communication. Upon the therapist's selection of a game mode, the board receives JSON-formatted information that is stored on a microSD card for processing.

The physical components consist of a board manufactured using FDM 3D printing with PETG filament, featuring a 4×2 grid of cells with integrated RFID readers and LED strips. A 4,000 mAh LiPo battery powers the board and includes an ESP32 microcontroller with WiFi connectivity, audio speaker, and micro-SD card storage. The stimuli system comprises physical Tangible Argument Objects, containing a bottom plastic block with RFID technology and a detachable and interchangeable top representation (featuring objects, pictures, or words), connected via velcro with the bottom part.

The software implementation utilizes a Nuxt.js web application for the therapist interface with authentication and patient management capabilities, supported by a Node.js server backend deployed through Docker containerization. A centralized database collects session metrics, including performance data and scheduling information, while a dedicated Node bridge server manages secure WebSocket communication between the board and web application. The ESP32 microcontroller is programmed in C++, enabling wireless communication and interactive feedback through colored LED responses and audio cues.

5 Experimental Study

The primary aim of our study was to evaluate the effectiveness of Moovy as an innovative tool for language rehabilitation and enhancement, based on the AAP (previously described in Section 3), which proposes that visually representing syntactic operations can be achieved through direct manipulation of stimuli. Thus, combining the mental representation of argumentative movement with its physical manifestation may enhance the comprehension of trained morphosyntactic constructions.

5.1 Research Questions

Since we aimed to investigate the efficacy of Moovy in facilitating the learning of targeted morphosyntactic constructs in children with DLD, we *hypothesized* that children undergoing training with Moovy will demonstrate a higher production of clitic pronouns and passive sentences compared to traditional methods.

The experimental design of our study was a 2×2 within-between design, with the factors *Group* (Experimental vs. Control) and *Evaluation* (results of tests on clitics and passive competence measured before vs. after the training). Both *Group(s)* underwent a 4-week training on *clitic pronouns*: in the *Experimental* group, the training was carried out using Moovy, and in the *Control* group, in traditional speech therapy sessions. Summarizing, our research question is: “Does training with Moovy enhance the production of clitic pronouns in children with DLD more effectively than traditional speech therapy methods?”

Additionally, since sentences with clitics and passive sentences share a similar argument movement, we aimed to assess whether the benefits of training on clitic pronouns generalize to passive sentences in order to assess the impact and strength of the AAP.

5.2 Target Group and Metrics

The study involved 30 Italian monolingual children with DLD, 12 girls and 18 boys, aged 5–8 years old ($M = 6.82$, $SD = 1.58$). They were recruited from clinical centers in Italy, where they were already undergoing clinical services. All children included in the study had received a diagnosis of DLD from neuro-psychiatric specialists based on standard inclusion and exclusion criteria [54]. We initially evaluated children’s grammatical abilities and nonverbal cognitive abilities. Grammatical abilities were evaluated by the “**Test di Comprensione Grammaticale per Bambini (TCGB)**” [18], or by the “**Batteria per la valutazione del linguaggio in bambini dai 4 ai 12 anni (BVL_4-12)**” [49]. Children were administered one of the two tests depending on the availability at the centers where the children were recruited and tested. Non-verbal cognitive abilities were evaluated using the standardized “Raven’s Coloured Progressive Matrices” test [58]. We administered a clitic production test provided by Arosio et al. [5], before and after the training. The test involves a computer-based elicitation task where children view a vignette described by a recorded voice. Subsequently, another vignette appears on the screen, and the voice asks a question that prompts the use of clitic pronouns in the answer. The test includes 3 familiarization trials and 10 experimental trials. Pre- and post-assessment tests differ in lexical materials [5]. Moreover, we also administered a passive sentence comprehension test by Volpato et al. [75] before and after the training. The test involves a sentence-image matching task. During the test, a recorded voice utters a sentence containing a passive construction, and the child is required to indicate the corresponding image from two alternatives appearing on the screen of the PC administering the task. The test includes five familiarization trials and 16 experimental trials. Pre- and post-assessment tests differ in lexical materials [75].

We equally divided the participants into two *Groups* (i.e., Control and Experimental), including 15 subjects each. The children in the Experimental group ($M = 6.18$, $SD = 2.13$) were matched with those in the Control group ($N = 15$, $M = 6.56$, $SD = 3.76$). All children had age-appropriate non-verbal IQ and scores on the Raven. Independent sample *t*-tests showed no significant difference between the raw scores of the Raven of the two groups ($t(28) = 0.840$, $p = 0.408$). The mean Raven score was 23.6 ($SD = 4.8$) in the Experimental group and 22.0 ($SD = 5.5$) in the Control group. While all children had a non-verbal IQ within normal limits, children with DLD scored 1.5 SD below the mean score for their age in the measures of grammatical abilities. Children with impairments in domains other than language or with a below-average score on Raven were excluded from the study.

5.3 Procedure

Before the study began, parental informed consent was obtained, and speech therapists underwent training on the materials and the functioning of Moovy. The entire procedure spanned 6 weeks, including 1 pre-assessment week, 4 training weeks, and 1 post-assessment week. During the first week, the recruited children underwent the assessment of verbal and nonverbal cognitive skills,

which included Raven's matrices, the TCGB/BVL, and tests assessing clitic pronoun production and passive sentence comprehension (previously described in Section 5.2). Additionally, this week, participants in the Experimental group underwent an initial familiarization session with Moovy, involving the warm-up activity. This activity aimed to help the child understand how the game functions. Over the subsequent 4 weeks, in the experimental group, the actual training on clitics occurred, with 4 sessions (1 per week), each lasting approximately 15 minutes. The training focuses exclusively on clitic pronouns. These sessions took place in a quiet room at the speech center, free from any additional materials to avoid distractions for the child.

In the Experimental group, during each training session, the speech therapist, using Moovy, presents the child with 16 pairs of sentences. Among these, 12 sentences include the targeted morphosyntactic construct with a primary focus on clitic pronouns, while the remaining 4 sentences follow the canonical subject–verb–object structure and are referred to as filler sentences. The inclusion of filler sentences aims to sustain the child's attention by introducing variations from the model, thereby minimizing potential biases related to the mechanical nature of the movement involved in understanding argumentative movement. It should be noted that only the tangible movements performed on the surface grid were recorded; no additional data—such as hand kinematics, hesitations, or correction gestures—were captured via video recording or other means. Similarly, the temporal coupling between vocal production and object manipulation (e.g., whether children verbalized sentences before, during, or after moving the tangibles) was not required for the child and therefore not annotated. This choice was deliberate, as requiring them to perform only the movement data allowed us to maintain experimental cleanliness and avoid introducing noise into the dependent variables.

Subsequently, the week following the completion of the training is dedicated to post-assessment. During this phase, the children undergo the same tests again to evaluate clitic pronoun production and passive sentence comprehension. This post-assessment allows for a comparison with the pre-tests to determine whether any changes have occurred as a result of the training.

In the Control group, children completed speech exercises focused on clitic pronouns following a traditional approach. The traditional training was scheduled for 15 minutes once a week for 4 weeks. For the remaining protocol, the children in the control group replicated the activities of the experimental group during the first and sixth weeks. Each session took place individually, with only one therapist present in a dedicated room.

Twelve therapists participated in this study, having been previously instructed and trained by researchers to perform the tasks consistently. Each therapist worked with both the experimental and the control group, ensuring a balanced distribution of therapist involvement across conditions. This arrangement was intentional, as it allows the therapist's effect to be considered neutralized, preventing any individual therapist's style or expertise from systematically biasing one condition over the other.

5.4 Test Scores and Data Analysis

Since the primary aim of our research was to evaluate Moovy's effectiveness as an innovative language training tool, children's responses to the clitic evaluation test were classified into five categories.

- *Target*: a response containing the clitic with the correct morphology (“*Lui la^(clitic pronoun)1 prende.*” [“*He it^(clitic pronoun) is taking.*”]).
- *Full DP*: all productions containing a full nominal constituent in post-verbal position instead of the clitic (“*Lui prende la farfalla.*” [“*He is catching the butterfly.*”]).

¹Third person singular feminine.

- *Omission*: responses in which the child omits the clitic (“*Lui prende.*” [“*He is taking.*”]);
- *Wrong Form*: sentences in which the clitic is produced but with the wrong morphology (“*Lui lo^(clitic pronoun)2 prende.*” [“*He it^(clitic pronoun) is taking.*”]).
- *Other*: any response that does not fit into the previous categories.

Responses in the passive sentence comprehension tests were categorized simply as correct or incorrect based on whether the child pointed to the picture corresponding to the sentence meaning or not.

All scores were analyzed using a **Generalized Linear Mixed Model (GLMM)** performed using the software R-studio after a careful cleaning process and data preparation.

5.5 Ethical Consideration

All study participants received prior notification regarding the study, including information about the study procedures, objectives, and data handling. They were recruited on a voluntary basis, and the studies were conducted in accordance with the Helsinki Declaration of 1975 (as revised in 2008). Ethical approval for the studies was obtained from the Ethical Committee of the University of Milano-Bicocca. The study includes preliminary tests to assess each child’s language proficiency as well as their verbal and non-verbal intelligence, in addition to evaluating a new learning tool. The experimental materials include plastic figurines commonly used and sold to children within the appropriate age range, figures for observational tasks, and writing materials.

6 Results

6.1 Descriptive Analysis

This section presents the descriptive outcomes of the results. For clitics production, children in the Control condition produced more target sentences than children in the Experimental condition during the pre-evaluation task. As we can observe in Figure 6 and Table 1, the Experimental group produced an average of 3.190 target sentences (SD = 2.977) out of 10, whereas the Control group produced an average of 4.533 (SD = 3.563). For both groups, this value represents a very low rate of production, as is typically observed in children with DLD. Instead of producing a *Target* sentence containing the clitic, children with DLD of both groups tended to omit the clitic and produce an ungrammatical sentence (*Omission*), produce a sentence with a post-verbal full nominal constituent instead of the clitic (*Full DP*), or produce *Other* responses (mainly including ungrammatical sentences containing non-finite verbs).

In the post-evaluation task, the production of *Target* sentences consistently increased in the Experimental group. They doubled the *Target* productions (M = 5.810, SD = 3.444) with a mean delta increase of 2.619 points. Conversely, in the Control group, the production of the *Target* sentences remained unchanged in the post-training evaluation (M = 5.133, SD = 3.420).

Concerning the comprehension of passive sentences, children in the Control group showed a lower comprehension rate than the experimental group at the pre-evaluation task. As shown in Figure 7 (and Table 1), the Experimental group achieved a mean comprehension score of M = 10.905 (SD = 2.982) points out of 16, whereas the Control group achieved a mean score of M = 11.133 (SD = 2.973) points at the pre-evaluation test. At the post-test, the average score of the Experimental group increased to M = 11.762 (SD = 3.404) points, with an average delta difference of M = 0.857 average score of the Control group remained substantially unchanged (M = 11.667, SD = 3.132—average delta 0.533, SD = 1.302).

²Third person singular masculine

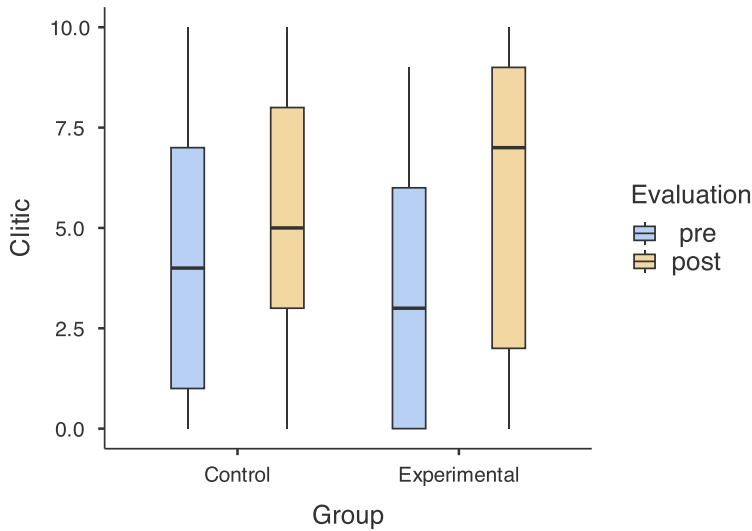


Fig. 6. Boxplot representing the production of clitic Target sentences in the pre- and post-tests for both the Control and Experimental groups.

Table 1. Descriptive Statistics of Results of the Pre- and Post-Tests with Delta Values in the Control and Experimental Groups

	Group	Clitic Pre	Clitic Post	Clitic Delta	Passive Pre	Passive Post	Passive Delta
Mean	Control	4.533	5.133	0.600	11.133	11.667	0.533
	Experimental	3.190	5.810	2.619	10.905	11.762	0.857
Median	Control	4	5	0	10	11	0
	Experimental	3	7	2	11	12	1
Standard deviation	Control	3.563	3.420	1.765	2.973	3.132	1.302
	Experimental	2.977	3.444	1.802	2.982	3.404	3.637
Minimum	Control	0	0	-1	7	6	-1
	Experimental	0	0	0	6	4	-12
Maximum	Control	10	10	6	16	16	4
	Experimental	9	10	6	16	16	6

6.2 Statistical Analysis

6.2.1 Analysis of Clitic Production. In the analysis of clitic productions, we excluded data from children who made fewer than two errors in the clitics pre-test in both the Control and Experimental groups. The threshold of two errors was identified based on previous studies investigating the sensitivity and specificity of the test when used to identify children with DLD (cf. Arosio et al. [5]). According to these studies, children with DLD typically make at least two errors at the clitic production task. Furthermore, we excluded data from individuals with a Raven score below 25°. In this initial analysis, four children were excluded—three for achieving excessively high clitic scores and one for having a notably low Raven score. Results from the standardized grammar proficiency tests (TCGB and BVL) were not included in the analysis, as the two tests yield different raw score scales. However, when raw scores were converted into z-scores relative to age-matched typically

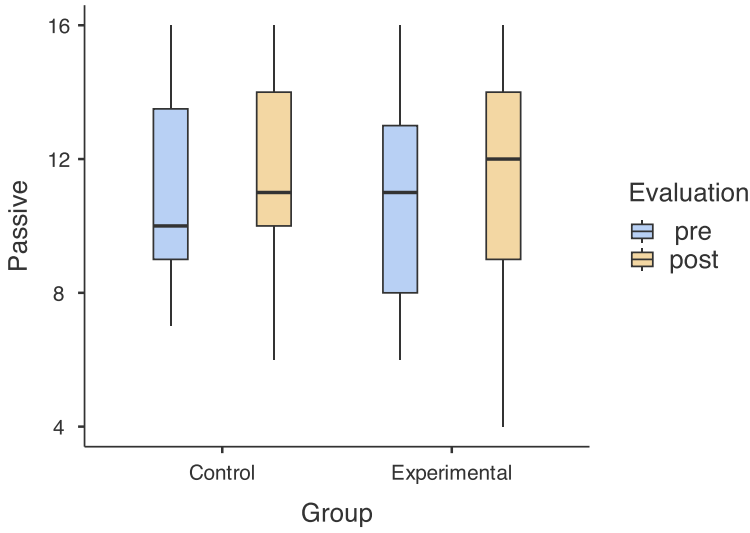


Fig. 7. Boxplot of passive comprehension in the pre- and post-test for the Control and Experimental groups.

developing peers, all children scored more than 1.5 standard deviations below the mean, confirming the presence of severe morphosyntactic deficits.

In our analysis, we employed a stepwise approach, an automated regression method particularly recommended for exploratory studies involving multiple variables, to estimate multivariate and/or GLMM. Since responses in each production category were coded as binomial dependent variables, we conducted three separate mixed model logistic regression analyses for the categories *Target*, *Full DP*, and *Other*. We did not consider *Omission* and *Wrong form* since there was no variation in the group answers. Our model included the interaction between Group (Experimental vs. Control) and Evaluation (pre vs. post) as a fixed factor, with random intercepts for subjects and items. This structure was identified as the maximal converging glmer model (lme4 package [9]) to describe *Target*, as determined by the buildmer function (buildmer package [73]). Therefore, we chose to apply this model structure to all the other dependent variables.

6.2.2 Analysis of Target Productions. This section reports the results of a GLMM analysis evaluating the effects of Group and Evaluation (as well as the interaction between them), on the *Target* variable.

Examining the target clitic productions, the interaction between Group and Evaluation significantly influenced the model ($\chi^2 = 15.764$, $p < 0.001$). Specifically, children in the Control group reported a significantly lower number of target productions than those in the Experimental group. Indeed, the *Target* productions in the children in the experimental group were significantly higher in the post-evaluation tasks with respect to the pre-evaluation task (pre vs. post: $\beta = 2.1239$, $p < 0.001$). This result suggests that the training program was effective in enhancing the production of clitics in DLD children over a period of 1 month of treatment. The analysis strongly suggests that both Group and Evaluation independently and interactively influence the target variable (reported in Table 2).

6.2.3 Analysis of Full DP Productions. We assessed the impact of Group and Evaluation (plus their interaction) on the *Full DP* variable using the GLMM. The Group did not significantly improve scores ($\chi^2 = 2.2016$, $p = 0.138$). A comparison was also made between incorporating Evaluation,

Table 2. Fixed Effects Estimates for the *Target* Productions

Fixed Effects	Estimate	Std. Error	z-Value	Pr(> z)
(Intercept)	-1.6644	0.5139	-3.239	<0.001
Group	0.7372	0.7277	1.013	0.311050
Evaluation	2.1239	0.3573	5.944	<0.001
Group $\times$ Evaluation	-1.8309	0.4716	-3.883	<0.001

Table 3. Fixed Effects Estimates for the *Full DP* Productions

Fixed Effects	Estimate	Std. Error	z-Value	Pr(> z)
(Intercept)	-1.8737	0.4207	-4.454	<0.001
Group	0.3832	0.5702	0.672	0.501557
Evaluation	-1.5296	0.4411	-3.468	<0.001
Group $\times$ Evaluation	1.1640	0.5589	2.083	0.037291

Table 4. Fixed Effects Estimates of the *Other* Productions

Fixed Effects	Estimate	Std. Error	z-Value	Pr(> z)
(Intercept)	-2.33849	0.55272	-4.231	<0.001
Group	0.01402	0.75832	0.018	0.98525
Evaluation	-1.46796	0.48764	-3.010	0.00261
Group $\times$ Evaluation	1.69945	0.62714	2.710	0.00673

demonstrating a significant improvement ($\chi^2 = 10.462$, $p < 0.001$). The interaction among the two variables also significantly improved performances ($\chi^2 = 6.6638$, $p = 0.0357$).

The model including the interaction (glmer.3) highlighted significant effects: the effect of Evaluation was significant ($\beta = -1.5296$, $p < 0.001$). The interaction Group $\times$ Evaluation) was also significant ($\beta = 1.1640$, $p = 0.037$), suggesting that the children in the Control group reported a significantly higher number of *Full DP* productions than those in the Experimental group.

Moreover, the analysis indicates that the factor Evaluation, both as a main effect and in interaction with Group, significantly affects the distribution of the *Full DP* variable. This result points to a complex interplay between group and evaluation, as reported in Table 3.

6.2.4 Analysis Other Productions. Finally, we analyzed the influence of Group and Evaluation on the *Other* variable. We first assessed the base effect with a first model and a model with the Group effect. The addition of the Group did not show a significant improvement ($\chi^2 = 0.9213$, $p = 0.3371$), so there was no difference between the Experimental and Control groups. The first model was compared with a model incorporating Evaluation, showing a marginally significant improvement ($\chi^2 = 2.9751$, $p = 0.0846$). A last comparison was made between the model with both Group and Evaluation interaction. The interaction model significantly improved ($\chi^2 = 8.778$, $df = 2$, $p = 0.0124$). The interaction model (glmer.3) revealed significant effects: Evaluation had a significant negative effect ($\beta = -1.46796$, $p = 0.0026$). The interaction term Group $\times$ Evaluation was significant ($\beta = 1.69945$, $p = 0.00673$), showing that children in the Control group made more production that was not classifiable as *Target* or *Full DP* than children in the Experimental group, as also reported in Table 4.

Table 5. Fixed Effects Estimates of the *Passive Sentences*

Fixed Effects	Estimate	Std. Error	z-Value	Pr(> z)
(Intercept)	1.09477	0.18264	5.99401	<0.001
Group	-0.07230	0.36051	-0.20055	0.84105
Evaluation	-0.23985	0.14023	-1.71044	0.08718
Group $\times$ Evaluation	-0.10472	0.28039	-0.37349	0.70878

6.2.5 Comprehension of Passive Sentences. While the accuracy of passive sentence comprehension was not the main focus of our study, we did measure it to explore the possibility of a transfer effect between different language structures. This transfer effect could be related to neural mechanisms such as priming and pre-activation of shared neural resources or short-term plasticity within a common neural network. However, it's important to note that our results for passive sentences did not reach statistical significance, and we lack sufficient evidence to suggest an observable effect. It is worth mentioning that passives and clitics, despite sharing a similar argument movement, have distinct structural differences. Clitics, in particular, involve more complexity as they override the verb. While we expect Moovy to improve clitic-related aspects, it may also have some impact on passives with pre-verbal objects. Nonetheless, it's essential to acknowledge that the Moovy study was carried out focusing on the training of clitics, thus our understanding of the results on passives remains limited. For brevity, the results and their discussion will not be expanded in this manuscript; while for completeness, in Table 5, the results of the statistical model are reported.

7 Discussion

Moovy has shown promise as an intervention tool for children with DLD. The findings indicate a significant improvement in the production of clitic pronouns among children with DLD who received intervention with Moovy (i.e., Experimental group) compared to those who did not use Moovy (i.e., Control group) over the course of 1 month.

Moreover, our results also indicate a reduction of errors during clitic production in children who participated in the Experimental condition. These findings highlight the potential efficacy of Moovy as an intervention tool for improving the linguistic performance of children with DLD, specifically in the domain of clitic production. However, further investigation with a larger, more thoroughly examined sample is essential for a deeper understanding of its impact. Considering the difficulty children with DLD face in acquiring complex morphosyntactic skills over a short period [46], our initial findings are encouraging. Given that speech therapy programs typically span several months/years and are intensive [24], the positive outcomes observed in our study are noteworthy. Additionally, our research illuminates the effectiveness of the AAP, which underpins Moovy. The modest advancements seen with Moovy in a short timeframe support the paradigm's theoretical basis that assumes that movement and language are intrinsically connected [69]. This connection is explored in Moovy and by broadening its prospective uses, ranging from virtual assistants to smart home technologies and robotics. With the evolution of technology, new research is expected to refine Moovy's functionalities, making significant contributions to the fields of Human-Computer Interaction. Nonetheless, caution is needed as this line of research is still in its early stages, and further work is needed to achieve a more comprehensive understanding. In particular, alternative explanations for the treatment's effects should be considered beyond the initial hypothesis. Additional research is therefore necessary to clarify the multiple factors contributing to the observed improvements and to provide a more robust and nuanced account of the treatment's effectiveness.

In fact, beyond the AAP mechanism, alternative explanations for the observed results merit consideration. Repeated exposure to clitic structures across four weekly sessions may have contributed independently of the motor-syntactic link. The sentence-pair structure of the task may also have introduced a priming effect, facilitating clitic production independently of the physical movement. Additionally, the interactive nature of Moovy—requiring children to actively listen, select, and physically place objects—may have prompted a degree of reflective attention to syntactic structure, engaging metacognitive monitoring of linguistic choices alongside, rather than separately from, the embodied mechanism. While therapist effects were controlled through cross-group assignment, the novelty of the interface may further have heightened children’s attention in ways not fully separable from the manipulation itself. Disentangling these factors would require experimental conditions that systematically isolate motor enactment from other features of the TUI interaction.

8 Conclusion

This article presented the *AAP*, a novel intervention paradigm for DLD, and *Moovy*, a TUI grounded in this paradigm. Designed as a tabletop game, Moovy serves as a tool for speech therapy, targeting children with DLD through brief yet intensive training sessions that aim to stimulate neural mechanisms involved in the processing of both motor and linguistic hierarchical structures [38].

This work makes two primary contributions. First, it provides a contribution in the TUIs arena through the development of Moovy, a gamified TUI. Moovy is the probe of a working system able to enhance children’s acquisition of complex morphosyntactic structures. Its interactive and engaging approach to language learning positions TUIs as a valuable asset in speech therapy. The integration of gaming elements into linguistic exercises offers a novel, more captivating method of language instruction, positively validated with a study involving $N=30$ participants over a 6-week period. In the domains of cognitive science and linguistics, Moovy signifies a paradigm shift, moving beyond traditional language therapy techniques. This tool represents a sophisticated blend of linguistic complexity and motor coordination, offering a nuanced approach to treating language disorders and facilitating language acquisition.

Second, our work provides a theoretical contribution to the understanding of the relationship between language and movement, reinforcing long-standing theories regarding the brain’s evolution and functioning. The theory positing the co-evolution of language and action, suggesting that certain linguistic constructs may have evolved from an ancestral motor system, finds support in our findings and confirms previous research in the field (cf. [3, 27]). These suggest a shared neural architecture for syntax processing across different domains. Thus, the AAP offers an innovative avenue for addressing DLD and enhancing language acquisition. By integrating the intricate interplay between language and the body, this paradigm redefines traditional language intervention methods.

Acknowledgement

The authors would like to thank all the participants, parents, and therapists involved in the study.

Author Statement

This submission presents the *AAP* and its evaluation with children with DLD.

Our prior work (cf. Beccaluva et al. [11]) provided only a preliminary introduction to the Moovy interface and focused exclusively on user experience evaluation with typically developing children. That work did not present the theoretical AAP framework, include participants with DLD, or conduct efficacy studies. The current submission makes substantial novel contributions: (i) full articulation of the AAP theoretical framework and its theoretical background, (ii) a rigorous 6-week

experimental study with 30 children with DLD, (iii) quantitative evidence of efficacy using GLMM analysis, and (iv) validation showing significant improvements in morphosyntactic skills.

This represents an entirely new contribution for ACM TOCHI, beyond the preliminary usability findings of our previous work. This manuscript has not been previously submitted to any other venue or journal and is not currently under consideration elsewhere.

References

- [1] Artemis Alexiadou and Edit Doron. 2012. The syntactic construction of two non-active voices: Passive and middle. *Journal of Linguistics* 48, 1 (2012), 1–34.
- [2] Alissa N. Antle, Min Fan, and Emily S. Cramer. 2015. PhonoBlocks: A tangible system for supporting dyslexic children learning to read. In *Proceedings of the 9th International Conference on Tangible, Embedded, and Embodied Interaction (TEI '15)*. ACM, New York, NY, 533–538. DOI: <https://doi.org/10.1145/2677199.2687897>
- [3] Susan C. Antón, Richard Potts, and Leslie C. Aiello. 2014. Evolution of early homo: An integrated biological perspective. *Science* 345, 6192 (2014), 1236828.
- [4] Fabrizio Arosio, Flavia Adani, and Maria Teresa Guasti. 2009. Grammatical features in the comprehension of Italian relative clauses by children. In *Merging Features: Computation, Interpretation, and Acquisition*. José M. Brucart, Anna Gavarró, and Jaume Solà (Eds.), Oxford University Press, 138–156.
- [5] Fabrizio Arosio, Chiara Branchini, Lina Barbieri, and Maria Teresa Guasti. 2014. Failure to produce direct object clitic pronouns as a clinical marker of SLI in school-aged Italian speaking children. *Clinical Linguistics & Phonetics* 28, 9 (2014), 639–663.
- [6] Fabrizio Arosio, Chiara Branchini, Matteo Forgiarini, Elena Roncaglione, Eleonora Carravieri, Emanuela Tenca, and M. Guasti. 2010. SLI children's weakness in morphosyntax and pragmatics. In *Proceedings of the 10th Tokyo Conference on Psycholinguistic*. Hituzi Syobo Press.
- [7] Fabrizio Arosio, Francesca Panzeri, Bruna Molteni, Santina Magazù, and M. Guasti. 2017. The comprehension of Italian relative clauses in poor readers and in children with specific language impairment. *Glossa* 2, 1 (2017), 9.
- [8] Fabrizio Arosio, Silvia Silleresi, and Maria Teresa Guasti. 2024. The production of relative clauses in Italian-speaking children with DLD. *Isogloss: Open Journal of Romance Linguistics* 10, 1 (2024), 1–23.
- [9] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2014. Fitting linear mixed-effects models using lme4. arXiv:1406.5823. Retrieved from <https://arxiv.org/abs/1406.5823>
- [10] Eleonora Beccaluva, Fabiano Riccardi, Mattia Gianotti, Jessica Barbieri, and Franca Garzotto. 2021. VIC—A tangible user interface to train memory skills in children with intellectual disability. *International Journal of Child-Computer Interaction* 1 (2021), 100376. DOI: <https://doi.org/10.1016/j.ijcci.2021.100376>
- [11] Eleonora Aida Beccaluva, Fabiano Riccardi, Giulia Bettelli, Lukasz Moskwa, Mariagiovanna Di Iorio, Francesco Riccardo Di Gioia, Eleonora Pasqua, Fabrizio Arosio, Maria Teresa Guasti, and Franca Garzotto. 2022. Grasping the grammar with Moovy: A tangible user interface to train linguistic skills in children. In *Proceedings of the CHI Conference on Human Factors in Computing Systems Extended Abstracts*, 1–7.
- [12] Adriana Belletti. 1999. Italian/Romance clitics: Structure and derivation. In *Clitics in the Languages of Europe*. Henk van Riemsdijk (Ed.), De Gruyter Mouton, 543–580.
- [13] Adriana Belletti and Chris Collins. 2020. *Smuggling in Syntax*. Oxford University Press.
- [14] Adriana Belletti and Maria Teresa Guasti. 2015. *The Acquisition of Italian: Morphosyntax and Its Interfaces in Different Modes of Acquisition*. John Benjamins Publishing Company
- [15] Giulia Bettelli, Maria Teresa Guasti, Paola Francesca Ajmone, Emanuela Tenca, and Fabrizio Arosio. 2023. The effect of a priming-based training on the production of object clitic pronouns in Italian speaking children with DLD. *Clinical Linguistics & Phonetics* 38 (2023), 642–675.
- [16] David R. Beukelman and Pat Mirenda. 2013. *Augmentative and Alternative Communication: Supporting Children and Adults with Complex Communication Needs* (4th ed.). Paul H. Brookes Publishing.
- [17] Umberta Bortolini, Barbara Arfé, Cristina M. Caselli, Luisa Degasperri, Patricia Deevy, and Laurence B. Leonard. 2006. Clinical markers for specific language impairment in Italian: The contribution of clitics and non-word repetition. *International Journal of Language & Communication Disorders* 41, 6 (2006), 695–712.
- [18] Anna Maria Chilosi and Paola Cipriani. 1995. *TCGB: Test di Comprensione Grammaticale per Bambini: Manuale*. Edizioni del Cerro.
- [19] Noppon Choosri, Chompoonut Pookao, Napatsayod Swangtrakul, and Anthony Atkin. 2017. Tangible interface game for stimulating child language cognitive skill. *IADIS International Journal on WWW/Internet* 15, 2 (2017), 17–31.
- [20] Alessandro Ciccale, Elisabetta Ambron, Angelika Lingnau, and Raffaella I. Rumiati. 2014. A kinematic analysis of age-related changes in grasping to use and grasping to move common objects. *Acta Psychologica* 151 (2014), 134–142.

- [21] Chris Collins. 2005. A smuggling approach to the passive in English. *Syntax* 8, 2 (Aug. 2005), 81–120. DOI : <https://doi.org/10.1111/j.1467-9612.2005.00076.x>
- [22] Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell, III, John C. Maxwell, and Annie E. Zaenen. 1995. *Formal Issues in Lexical-Functional Grammar*. Center for the Study of Language (CSLI), Stanford.
- [23] Chris Dollaghan and Thomas F. Campbell. 1998. Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research* 41, 5 (Oct. 1998), 1136–1146. DOI : <https://doi.org/10.1044/jslhr.4105.1136>
- [24] Susan H. Ebbels, Lisa Wright, Sally Brockbank, Caroline Godfrey, Catherine Harris, Hannah Leniston, Kate Neary, Hilary Nicoll, Lucy Nicoll, Jackie Scott, et al. 2017. Effectiveness of 1:1 speech and language therapy for older children with (developmental) language disorder. *International Journal of Language & Communication Disorders* 52, 4 (2017), 528–539.
- [25] Luciano Fadiga and Laila Craighero. 2006. Hand actions and speech representation in Broca’s area. *Cortex* 42, 4 (2006), 486–490.
- [26] Luciano Fadiga, Laila Craighero, and Alice Roy. 2006. *Broca’s Region: A Speech Area*. Oxford University Press Oxford, Oxford. 137–152.
- [27] W. Tecumseh Fitch. 2011. Unity and diversity in human language. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 366, 1563 (2011), 376–388.
- [28] Raffaella Folli and Heidi Harley. 2013. The syntax of argument structure: Evidence from Italian complex predicates. *Journal of Linguistics* 49, 1 (2013), 93–125.
- [29] Charles Randy Gallistel. 2011. Prelinguistic thought. *Language Learning and Development* 7, 4 (2011), 253–262.
- [30] Franca Garzotto and Manuel Bordogna. 2010. Paper-based multimedia interaction as learning tool for disabled children. In *Proceedings of the 9th International Conference on Interaction Design and Children*, 79–88.
- [31] Franca Garzotto and Roberto Gonella. 2011. Children’s co-design and inclusive education. In *Proceedings of the 10th International Conference on Interaction Design and Children (IDC ’11)*. ACM, New York, NY, 260–263. DOI : <https://doi.org/10.1145/1999030.1999077>
- [32] Mirko Gelsomini, Micol Spitale, and Franca Garzotto. 2021. Phygital interfaces for people with intellectual disability: An exploratory study at a social care center. *Multimedia Tools and Applications* 80, 26 (2021), 34843–34874.
- [33] Arthur M. Glenberg. 2011. How reading comprehension is embodied and why that matters. *International Electronic Journal of Elementary Education* 4, 1 (2011), 5–18.
- [34] Arthur M. Glenberg, Andrew B. Goldberg, and Xiaojin Zhu. 2011. Improving early reading comprehension using embodied CAI. *Instructional Science* 39, 1 (2011), 27–39.
- [35] Theres Grüter and Martha Crago. 2012. Object clitics and their omission in child L2 French: The contributions of processing limitations and L1 transfer. *Bilingualism: Language and Cognition* 15, 3 (2012), 531–549.
- [36] Maria Teresa Guasti. 1993. Verb syntax in Italian child grammar: Finite and nonfinite verbs. *Language Acquisition* 3, 1 (1993), 1–40.
- [37] Bart Hengeveld, Caroline Hummels, Hans van Balkom, Riny Voort, and Jan de Moor. 2013. Wrapping up LinguaBytes, for now. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction (TEI ’13)*. ACM, New York, NY, 237–244. DOI : <https://doi.org/10.1145/2460625.2460664>
- [38] Satomi Higuchi, Thierry Chaminade, Hiroshi Imamizu, and Mitsuo Kawato. 2009. Shared neural correlates for language and tool use in Broca’s area. *Neuroreport* 20, 15 (2009), 1376–1381.
- [39] Hiroshi Ishii and Brygg Ullmer. 1997. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI ’97)*. ACM, New York, NY, 234–241. DOI : <https://doi.org/10.1145/258549.258715>
- [40] Ray Jackendoff and Steven Pinker. 2005. The nature of the language faculty and its implications for evolution of language (reply to Fitch, Hauser, and Chomsky). *Cognition* 97, 2 (2005), 211–225.
- [41] Janio Jadan-Guerrero, Javier Jaen, Maria A. Carpio, and Luis A. Guerrero. 2015. Kiteracy: A kit of tangible objects to strengthen literacy skills in children with down syndrome. In *Proceedings of the 14th International Conference on Interaction Design and Children (IDC ’15)*. ACM, New York, NY, 315–318. DOI : <https://doi.org/10.1145/2771839.2771905>
- [42] Celia Jakubowicz, Lea Nash, Catherine Rigaut, and Christophe-Loic Gerard. 1998. Determiners and clitic pronouns in French-speaking children with SLI. *Language Acquisition* 7, 2–4 (1998), 113–160.
- [43] Jonathan Kleiman, Michael Pope, and Paulo Blikstein. 2013. RoyoBlocks: An exploration in tangible literacy learning. In *Proceedings of the 12th International Conference on Interaction Design and Children (IDC ’13)*. ACM, New York, NY, 543–546. DOI : <https://doi.org/10.1145/2485760.2485861>
- [44] Alistair Knott. 2012. *Sensorimotor Cognition and Natural Language Syntax*. MIT Press, Boston.
- [45] Kathryn Kohnert, Jennie Windsor, Kerry D. Ebert, Haleigh Larrabee, and Amy M. Walk. 2020. Virtual reality treatment (VRT) as a new tool for bilingual children with developmental language disorders. *Journal of Communication Disorders* 87 (2020), 106010.

- [46] James Law, Zoe Garrett, and Chad Nye. 2003. Speech and language therapy intervention for children with primary speech and language delay or disorder: Cochrane developmental, psychosocial and learning problems group. *Cochrane Database of Systematic Reviews* 3, 02 (2003), CD004110. DOI: <https://doi.org/10.1002/14651858.CD004110>
- [47] Laurence B. Leonard. 2014. *Children with Specific Language Impairment*. MIT Press, Boston.
- [48] Giovanni Manetti. 2013. *Principio Era il Segno. Momenti di Storia Della Semiotica Nell'antichità Classica*. Bompiani.
- [49] Andrea Marini, Luigi Marotta, Sara Bulgheroni, and Franco Fabbro. 2015. *Batteria per la Valutazione del Linguaggio in Bambini dai 4 ai 12 Anni (BVL_4-12)*. Guinti O.S.
- [50] Lotte Meteyard, Sara Rodriguez Cuadrado, Bahador Bahrami, and Gabriella Vigliocco. 2012. Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex* 48, 7 (2012), 788–804.
- [51] Timothy Neate, Abi Roper, Stephanie Wilson, Jane Marshall, and Madeline Cruice. 2020. CreaTable content and tangible interaction in aphasia. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. ACM, New York, NY, 1–14. DOI: <https://doi.org/10.1145/3313831.3376490>
- [52] William Noble and Iain Davidson. 1996. *Human Evolution, Language and Mind: A Psychological and Archaeological Inquiry*. CUP Archive, Cambridge University Press.
- [53] Stefan Oppl and Chris Stry. 2011. Towards informed metaphor selection for TUIs. In *Proceedings of the 3rd ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, 247–252.
- [54] World Health Organization. 2004. *ICD-10: International Statistical Classification of Diseases and Related Health Problems: Tenth Revision*. World Health Organization.
- [55] K. J. Reddy. 2025. Speech and language rehabilitation technologies. In *Innovations in Neurocognitive Rehabilitation: Harnessing Technology for Effective Therapy*. Springer Nature Switzerland, Cham, 211–229.
- [56] Philippe Prévost. 2006. The phenomenon of object omission in child L2 French. *Bilingualism: Language and Cognition* 9, 3 (2006), 263–280.
- [57] Friedemann Pulvermüller and Luciano Fadiga. 2010. Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews. Neuroscience* 11, 5 (2010), 351–360.
- [58] John C. Raven and John Hugh Court. 1998. *Raven's Progressive Matrices and Vocabulary Scales*, Vol. 759. Oxford Psychologists Press, Oxford.
- [59] Rasulov Ravshankhodja. 2020. Word valence and syntactic relationship. *International Journal of Psychosocial Rehabilitation* 24, 2 (2020), 295–300.
- [60] Alice C. Roy, Aurore Curie, Tatjana Nazir, Yves Paulignan, Vincent Des Portes, Pierre Fourneret, and Viviane Deprez. 2013. Syntax at hand: Common syntactic structures for actions and language. *PLoS One* 8, 8 (2013), e72677.
- [61] Cinzia Russi. 2008. *Italian Clitics: An Empirical Study*. Mouton de Gruyter.
- [62] Joaquim Santos, Luis M. T. Jesus, Mario Vairinhos, and Jonathan Rodriguez. 2019. Tangible user interface for homework for children with communication disorders. In *Proceedings of the 2019 14th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, 1–6.
- [63] B. Schneider, P. Jermann, G. Zufferey, and P. Dillenbourg. 2011. Benefits of a tangible interface for collaborative learning and interaction. *IEEE Transactions on Learning Technologies* 4, 3 (2011), 222–232. DOI: <https://doi.org/10.1109/TLT.2010.36>
- [64] Lawrence Shapiro. 2014. *The Routledge Handbook of Embodied Cognition*. Routledge, UK.
- [65] Lawrence Shapiro. 2019. *Embodied Cognition*. Routledge.
- [66] Marlaine C. Smith. 2019. *Sensory Integration: Theory and Practice*. FA Davis.
- [67] C. Sylla, I. S. P. Pereira, C. P. Coutinho, and P. Branco. 2016. Digital manipulatives as scaffolds for preschoolers' language development. *IEEE Transactions on Emerging Topics in Computing* 4, 3 (2016), 439–449. DOI: <https://doi.org/10.1109/TETC.2015.2500099>
- [68] Victoria Tam, Mirko Gelsomini, and Franca Garzotto. 2017. Polipo: A tangible toy for children with neurodevelopmental disorders. In *Proceedings of the 11th International Conference on Tangible, Embedded, and Embodied Interaction*, 11–20.
- [69] Simon Thibault, Raphaël Py, Angelo Mattia Gervasi, Romeo Salemmé, Eric Koun, Martin Lövdén, Véronique Boulenger, Alice C. Roy, and Claudio Brozzoli. 2021. Tool use and language share syntactic processes and neural patterns in the basal ganglia. *Science* 374, 6569 (2021), eabe0874.
- [70] Heather K. J. Van der Lely. 2005. Domain-specific cognitive systems: Insight from Grammatical-SLI. *Trends in Cognitive Sciences* 9, 2 (2005), 53–59.
- [71] Maria Vender, Maria Garraffa, Antonella Sorace, and Maria Teresa Guasti. 2016. How early L2 children perform on Italian clinical markers of SLI: A study of clitic production and nonword repetition. *Clinical Linguistics & Phonetics* 30, 2 (2016), 150–169.
- [72] Maria Vender, Shenai Hu, Federica Mantione, Denis Delfitto, and Chiara Melloni. 2018. The production of clitic pronouns: A study on bilingual and monolingual dyslexic children. *Frontiers in Psychology* 9 (2018), 2301.
- [73] C. C. Voeten. 2022. Buildmer: Stepwise Elimination and Term Reordering for Mixed-Effects Regression. Retrieved from <https://CRAN.R-project.org/package=buildmer>

- [74] Francesca Volpato and Giulia Bozzolan. 2017. Explicit teaching of syntactic movement in passive sentences and relative clauses: The case of a Romanian/Italian sequential bilingual child. *Annali di Ca' Foscari. Serie Occidentale* 51 (2017), 357–381.
- [75] Francesca Volpato, Laura Verin, and Anna Cardinaletti. 2016. The comprehension and production of verbal passives by Italian preschool-age children. *Applied Psycholinguistics* 37, 4 (2016), 901–931.
- [76] Stefano Volpato, Julia Daragjati, Matteo Simonato, Andrea Fontana, Luigi Ferrucci, Alberto Pilotto, and MPI Tri Veneto Study Group. 2016. Change in the multidimensional prognostic index score during hospitalization in older patients. *Rejuvenation Research* 19, 3 (2016), 244–251.
- [77] Yuyang Wang, Yali Chen, Shujuan Lin, and Jianxin Wang. 2018. A mobile application to improve Chinese children's vocabulary: An efficacy and usability study. *Journal of Educational Computing Research* 56, 7 (2018), 1118–1137. DOI : <https://doi.org/10.1177/0735633117745901>
- [78] Yuyang Wang, Huijun Li, Xiaodong Li, and Tingting Ye. 2021. Effects of speech recognition technology on speech intelligibility in patients with cleft lip and palate: A randomized controlled trial. *Journal of Cranio-Maxillofacial Surgery* 49, 3 (2021), 213–219. DOI : <https://doi.org/10.1016/j.jcms.2020.12.016>
- [79] Dieter Wanner. 1987. Clitic pronouns in Italian: A linguistic guide. *Italica* 64, 3 (1987), 410–442.
- [80] Margaret Wilson. 2002. Six views of embodied cognition. *Psychonomic Bulletin & Review* 9, 4 (2002), 625–636.
- [81] Fangtian Ying, Pinhao Wang, Yuping Zou, Xinglin Zheng, Muling Huang, and Cheng Yao. 2021. TangiLetter: An interactive vocabulary system for children to learn words about wild animals. In *Distributed, Ambient and Pervasive Interactions*. Norbert Streitz and Shin'ichi Konomi (Eds.), Springer International Publishing, Cham, 287–298.

Appendix

A Argument Structure of Sentences, Clitics Pronouns, and Passive Sentences

In linguistics, a sentence is a maximal syntactic unit that is structurally complete according to the grammar of a language and can function independently as an utterance. Within sentences, individual words are organized into hierarchically structured sequences, forming units known as *phrases*. Phrases are linguistic structures that display unitary syntactic behavior and may consist of a single word or a combination of several elements [59]. Phrases can occupy argument positions and serve as *arguments* for a verb, to complete its semantic or grammatical structure [1]. For instance, consider the sentence “*The nice guy sends the letter.*” In this sentence, the phrases “*The nice guy*” and “*the letter*” are arguments of the verb “*send*.” The verb “*send*” is a transitive verb, and it requires two arguments, an *internal* one and an *external* one. In Italian, as in English, the typical linear order between the verb and its arguments in a finite sentence is *subject–verb–object*, where the subject position in the sentence is associated with the *external* verb argument and the object position with the *internal* verb argument [22]. In our example, the phrase “*The nice guy*” functions as the external argument of the verb, while “*the letter*” serves as the internal argument. However, both in Italian and in English, verb arguments can also appear in uncanonical orders, depending on some grammatical requirements. For instance, since the sentence subject position should be occupied by a discourse topic, when the discourse topic is the internal argument of the verb, the internal argument is associated with the sentence subject position, giving rise to the uncanonical word order found in passive sentences. When we talk about an *argument movement*, we refer to the displacement of a verb argument to a position different from its canonical one, as described above. In Italian, in addition to passive sentences, sentences containing DO clitic pronouns present a non-canonical order of the verb arguments. Object clitic pronouns are internal verb arguments realized by monosyllabic morphemes denoting discourse topics “*La maestra scrive una lettera e il ragazzo carino la^(clitic pronoun) spedisce.*” [“*The teacher writes a letter and the nice guy it^(clitic pronoun) sends.*”]. According to Belletti [12], the object clitic pronouns originate right to the verb, in its internal argument position, and later move to its left in an uncanonical position where discourse topics surface. Both finite sentences with DO clitic pronouns and passive sentences share a leftward movement of the internal verb argument to an uncanonical position. The leftward argument movement leads to a transformation in the linear representation of the sentence. The morphosyntactic operations necessary for these structural

changes are complex and demanding, requiring the child to have morphosyntactic competence, including understanding argument properties and functional procedures for constructing syntactic chains, in order to execute these operations. As mentioned earlier, children with DLD encounter the most challenges in the morphosyntactic domain, particularly with specific structures like clitic pronouns and passive sentences [15].

A.1 Clitic Pronouns

In Italian, clitic pronouns are morphemes that serve specific grammatical functions. DO clitic pronouns possess some complexity, making them particularly challenging to acquire [5, 71]. Phonologically, clitics are weak monosyllabic morphemes that are not independent, i.e., they cannot exist without a verb to refer to. Syntactically, their position changes on the basis of the finiteness of the verb [6]. If the verb is finite, clitics are placed in the pre-verbal position, preceding the predicate and resulting in a non-canonical subject–object–verb structure, for example, “*Il ragazzo carino la^(clitic pronoun) spedisce.*” [“*The nice guy it^(clitic pronoun) sends.*”]. In this case, they are referred to as proclitics. Conversely, if the verb is non-finite, clitics are positioned post-verbally, following the predicate, and the sentence maintains a canonical subject-verb-object order, “*La maestra dice al ragazzo carino di spedirla^(clitic pronoun).*” [“*The teacher says to the nice guy to sendit^(clitic pronoun).*”]; in this case, they are called enclitics [5].

At the morphosyntactic level, clitics in Italian are marked for gender, number, and person, resulting in specific forms. For instance, “*la*” is third person singular feminine, “*lo*” is third person singular masculine, “*le*” is third person plural feminine, and “*li*” is third person plural masculine [79]. Additionally, when the tense of the sentence is in the past, the verb must agree in gender and number with the clitic. In these cases, the contracted form is typically used. Furthermore, at the pragmatic level, a clitic can only refer to a previously introduced discourse topic. Consequently, using a DO clitic pronoun correctly requires the speaker to recognize the specific discourse topic to which it refers [61]. Italian-speaking children with typical language development commonly begin using DO clitics around the age of 2 years, mirroring adult patterns without making misplacements or substitutions with other pronouns. Instead, they occasionally replace the pronoun with a full nominal phrase (“*La nonna lava la frutta.*” [“*Grandma washes the fruit.*”] instead of “*La nonna la^(clitic pronoun) lava.*” [“*The grandmother washes it^(clitic pronoun).*”]), or they may omit the pronoun altogether (“*La nonna lava.*” [“*Grandma washes.*”] instead of “*La nonna la^(clitic pronoun) lava.*” [“*The grandmother washes it^(clitic pronoun).*”]). Moreover, these children demonstrate the ability to match the gender and number of the past participle verb with that of the clitic in sentences describing recent events. However, until around the age of 4 years, they experience a phase characterized by occasional omission of pronouns in contexts where it is expected [72].

The omission of DO clitics is a good clinical marker of DLD in Italian-speaking children. Specifically, children with DLD tend to omit clitics at the ages of 4–6 years [17], or they substitute it with a full nominal constituent at school ages [5]. Concerning the use of **Reflexive Clitics (RE)**, such as the clitic “*si*” in the sentence “*Il bambino si^(clitic pronoun) lava.*” [“*The child himself^(clitic pronoun) washes.*”], research indicates that both French- and Italian-speaking children exhibit a lower rate of omission for RE clitics compared to DO clitics [5, 42].

While the above-mentioned studies have focused on preschool-aged children, it is crucial to note that language development disorders are long-lasting conditions that often persist into adulthood. Some research has focused on examining the production of DO clitics in school-age children. The results of the study reported in Jakubowicz et al. [42] indicate that even beyond the age of 5 and up to at least age 10, children continue to struggle with the production of clitic pronouns. However, there is a notable distinction in their presentation: older children do not simply omit the internal

argument to the verb, resulting in ungrammatical sentences, as younger children do. Instead, they tend to insert a complete lexical nominal constituent in place of the pronoun within the sentence, positioned post-verbally. This pragmatic choice is inappropriate but grammatically correct. An example, resuming the above introduced sentences, if we consider “*Prima il bambino prende la mela. Poi lui la^(clitic pronoun) mangia.*” [“*First, the child takes the apple. Then he eats it^(clitic pronoun).*”], a form with a repetition of the lexical nominal constituent (in bold in the example) would be “*Prima il bambino prende la mela. Poi lui mangia **la mela.***” [“*First, the child takes the apple. Then he eats **the apple.***”].

Based on the findings from these studies, the ability to produce DO clitics emerges as a reliable marker for distinguishing between children with DLD and typically developing children of the same age. This distinction exhibits high levels of specificity and sensitivity, where sensitivity refers to the accuracy of identifying DLD subjects as true cases of DLD and specificity refers to the accuracy of identifying typically developing subjects as true cases of typical development. Similar challenges were identified at the cross-linguistic level. To explain the underlying causes of this difficulty, two primary hypotheses have been proposed. The first suggests that the complexity of producing clitics arises from the fact that these morphemes are weak and monosyllabic, making them phonologically insubstantial and consequently challenging to acquire [17]. However, this hypothesis is not supported by the observation that RE, which are also phonologically weak, are acquired and utilized correctly by children with DLD. An alternative explanation, put forth by Prevost, attributes the difficulties in mastering clitics to processing problems, specifically inadequate cognitive resources [56]. This hypothesis finds support in Grüter and Crago [35] study, which revealed a significant correlation between working memory and pronoun omission. Individuals with lower working memory capacity produced fewer clitics.

A.2 Passive Sentences

Passive sentences are intricate syntactic structures characterized by a word order different from the canonical one, resulting from a reorganization of grammatical functions. In this construction, the DO or complement of a transitive verb becomes the subject of the sentence. At the same time, the verb undergoes morphological changes, forming a structure with the auxiliary verb “to be” (or “venire” in Italian—that is literally translated in “to come” in English) followed by the past participle of the main verb. This composition is conjugated according to the number and gender of the newly established subject. Additionally, the agent responsible for the action may or may not be explicitly expressed through an adjunct phrase (a by-phrase), indicating the person or entity performing the action. In Italian, this element is commonly referred to as the *complemento d’agente* [complement of an agent]. This transformation converts an active sentence like “*The mother feeds the child.*” into its passive counterpart, i.e., “*The child is fed by the mother.*” [76].

Similar to clitic pronouns, the construction of passive sentences also involves a process of argument movement. Specifically, the internal argument of the verb (e.g., the child) moves to a pre-verbal position, while the external argument (e.g., the mother) is relegated to a post-verbal position within the sentence. However, the complexity of passive constructions surpasses that of clitics due to the initial blocking of object displacement into a pre-verbal position by the subject [74]. Belletti and Collins [13] introduced the *smuggling hypothesis* to propose a model for the derivation of passive sentences. According to this theory, the formation of a passive sentence occurs in two distinct steps: first, the verb and its object move together, bypassing the external argument (the mother). Since no argument occupies the pre-verbal position, the internal argument can then be raised to the left periphery of the sentence, facilitating the construction of the passive, as exemplified in “*The child is fed by the mother.*” [21].

The acquisition of the passive form in Italian remains a topic of debate, with studies conducted so far yielding conflicting results. The central question revolves around whether passive sentences are accurately represented from the early stages of language acquisition or if they are acquired at a later developmental stage.

Two recent studies assert that children are capable of having a correct representation of passive sentences right from the initial stages of learning Italian [48, 76]. Manetti [48] research reveals that Italian children between the ages of 3–6 and 4–6 can already produce passive sentences, both with “*essere*” [“*to be*”] and “*venire*” [“*to come*”] auxiliaries, when exposed to syntactic priming. Syntactic priming is a phenomenon where the processing of a specific language structure is facilitated when the same structure has been presented shortly before. In this context, the presentation of a passive prime aids the processing of the passive construction, demonstrating the ability of preschool children to grasp the syntactic structure in question.

The study conducted by Volpato et al. [76] investigated the comprehension and production of passive sentences in Italian-speaking children aged 3–6 years. The study compared three variables: the verb class (action verbs vs. non-action verbs), the presence or absence of a by-phrase, and the use of auxiliary verbs. Results suggest that children’s understanding of passive sentences precedes production. The results from the comprehension tasks indicated that young children possess knowledge of verbal passives from the earliest stages of language acquisition.

Younger children may face challenges in understanding passives with agent specification due to potential memory overload and their more limited cognitive resources compared to older children. The study replicated previous findings regarding the distinction between action and non-action verbs. It reaffirmed that action passives are better understood in all conditions (whether using the auxiliary verbs, with or without an accessory phrase). This suggests that the difference in comprehension does not stem from sentence ambiguity but rather from a greater difficulty in representing non-action verbs. Production results were consistent with comprehension findings: children were able to produce passive sentences using both “*essere*” [“*to be*”] and “*venire*” [“*to come*”] auxiliaries, with and without by-phrases. In conclusion, typically developing children between 3 and 6 years demonstrate the ability to comprehend and produce passive syntactic structures from the early stages of language acquisition.

Received 3 November 2025; revised 8 May 2026; accepted 4 June 2026