

# LingoLift: Supporting Educators in Personalized Oral Language Teaching for Autistic Children through Content Generation

Jiawen Zhang  
 jzhang731@connect.hkust-gz.edu.cn  
 The Hong Kong University of Science and Technology (Guangzhou)  
 Guangzhou, China

Dongyijie Primo Pan  
 dpan750@connect.hkust-gz.edu.cn  
 The Hong Kong University of Science and Technology (Guangzhou)  
 Guangzhou, China

Li Wang  
 nanshaqihui@outlook.com  
 Nansha Qihui School  
 Guangzhou, China

Pan Hui  
 panhui@hkust-gz.edu.cn  
 The Hong Kong University of Science and Technology (Guangzhou)  
 Guangzhou, China  
 The Hong Kong University of Science and Technology  
 Hong Kong, China

Xin Tong\*  
 Xint@hkust-gz.edu.cn  
 The Hong Kong University of Science and Technology (Guangzhou)  
 Guangzhou, China  
 The Hong Kong University of Science and Technology  
 Hong Kong, China

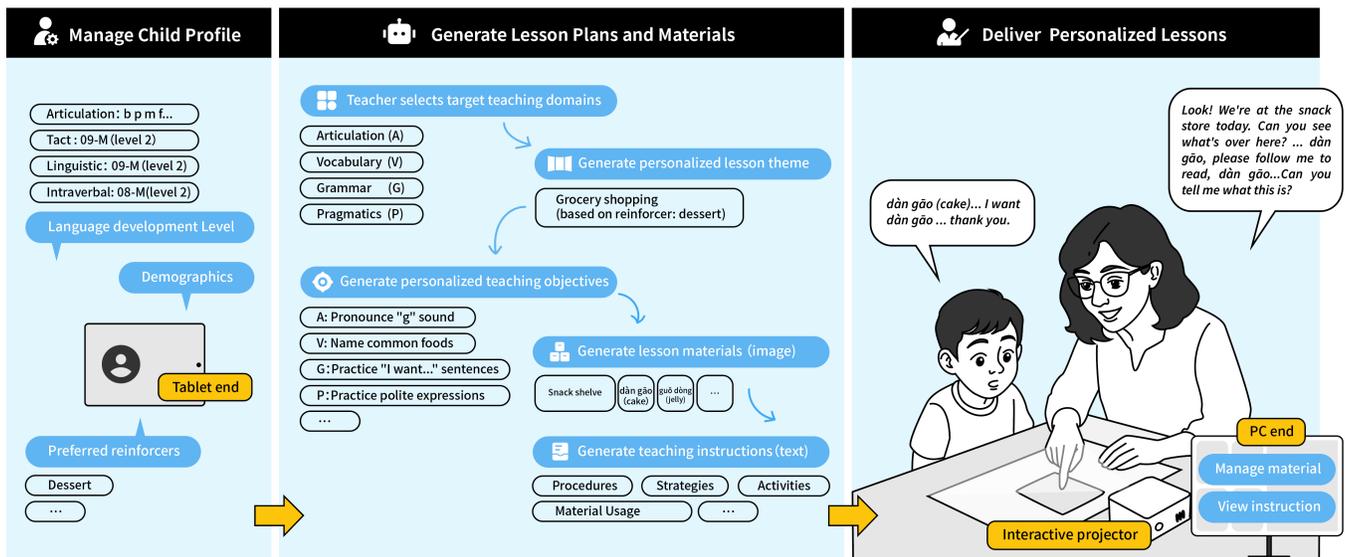


Figure 1: Teacher’s usage flow of LingoLift: (1) Manage Profile: adjust children’s language ability and learning interests (manual input required for initial setup, with automatic updates based on learning progress); (2) Prepare Lesson: LingoLift generates personalized learning plans and materials tailored to child’s capabilities and interests; (3) Deliver Lesson: Seamless integration with projection systems enabling direct use of generated content for instruction.

## Abstract

Autistic children exhibit heterogeneous oral language impairments, necessitating educators to implement personalized teaching content.

\*Corresponding author.



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

CHI '26, Barcelona, Spain

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

ACM ISBN 979-8-4007-2278-3/2026/04

<https://doi.org/10.1145/3772318.3790284>

However, preparing personalized materials remains time-intensive and difficult to maintain coherence, while generative AI’s recent advances in creating customized content show potential to support this process. We first conducted video analysis from educators’ one-on-one classes with autistic students and conducted interviews with therapists to understand their challenges in current teaching practices. Then, we developed a generative AI-empowered prototype, LingoLift, which supports educators to create interest-based, ability-adapted, and coherent teaching materials according to children’s profiles. Finally, we conducted a three-week deployment study with 10 educator-student dyads completing 30 lessons with

LingoLift in a specialized education school. Results showed that LingoLift significantly improved lesson preparation efficiency, reduced educators' workload, and enabled children to achieve positive learning outcomes. We observed educators' adaptive extensions and innovations, revealing insights into design considerations and future opportunities for AI-assisted inclusive education.

## CCS Concepts

• **Applied computing** → **Computer-assisted instruction**; • **Human-centered computing** → **Interactive systems and tools**; **Accessibility**; **User interface management systems**.

## Keywords

Autism Spectrum Disorder, Personalized Learning, Generative AI, Oral Language Learning, Human-AI Collaboration

### ACM Reference Format:

Jiawen Zhang, Dongyijie Primo Pan, Li Wang, Pan Hui, and Xin Tong. 2026. LingoLift: Supporting Educators in Personalized Oral Language Teaching for Autistic Children through Content Generation. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*, April 13–17, 2026, Barcelona, Spain. ACM, New York, NY, USA, 25 pages. <https://doi.org/10.1145/3772318.3790284>

## 1 Introduction

Autism Spectrum Disorder (ASD) is characterized by impairments or delays in oral language development [17, 28], with individuals exhibiting substantial heterogeneity [38]. This heterogeneity manifests in broad differences in oral language abilities (ranging from producing single syllables to constructing complex sentences [79]), diverse impairment patterns (including limited spontaneous communication [30, 57], lexical-intentional dissociation [70], and rigid syntactic structures [1]), as well as co-occurring conditions such as attentional deficits and motivational disorders [61, 65, 93]. Early oral language development significantly predicts long-term quality of life and reduces social isolation risks among autistic children [56, 72].

Given the multidimensional heterogeneity of autistic children, personalized teaching content design requires integrating hierarchical goal setting and motivational stimulation based on individual characteristics [35, 69]. Educators play a crucial role in designing and implementing such teaching practices [32, 85]. They develop and continuously refine teaching plans based on their familiarity with children [84], while demonstrating exceptional flexibility in addressing diverse learning challenges faced by autistic children [35]. However, this work is typically time-consuming [33, 83] and lacks appropriate resources [6, 52], constraining teachers' capacity to meet individual needs [81].

Recent HCI research has supported language learning for autistic children through VR immersive environments [11, 49], serious games for engagement [13, 43, 44], and AR/mobile applications for real-world generalization [18, 41, 62, 71, 96, 101]. However, these studies treat children as independent users, with limited educator involvement [23, 71]. Some research even emphasizes creating technology to completely replace practitioners [11], overestimating children's digital literacy and judgment [39] while failing to leverage educators' professional abilities in flexible adaptation

and meaningful social interaction [32, 59, 76], limiting learning outcomes [33]. Our emphasis on educator involvement aligns with recent work by Shin et al. [88], Porayska-Pomsta et al. [74], and Elbeledy et al. [19], who similarly highlight the importance of including educators in technology-mediated interactions for better observation, richer learning experiences, and more effective support. Additionally, prior work in technology-supported language learning for autistic children predominantly focuses on specific vocabulary learning [18, 25, 41, 62, 96, 101], inadequately addressing children's diverse oral language needs including articulation, grammar, and conversation [89, 99], limiting personalized applicability [35, 69]. Comparing with previous research, we propose educator-centered technological support [84, 103] addressing broader oral language learning needs [73, 94] for enhanced personalized teaching outcomes [46, 54].

Recent advances in generative AI have significantly improved learning outcomes for autistic children by creating customized content that matches cognitive levels [55, 97], integrates personal interests [60, 108], and provides real-time responses [51, 52, 91, 108]. These technologies help educators rapidly construct lesson plans and obtain teaching materials [20, 24, 63, 77, 78, 81], providing new possibilities for supporting oral language educators' teaching practices. Therefore, we aim to further explore:

**RQ1:** How can we design and implement AI tools to support educators in providing personalized oral language teaching for autistic children?

**RQ2:** How do AI generation capabilities change and support educators' personalized teaching practices?

We first conducted a formative study to understand current personalized teaching challenges and identify AI opportunities, including observations of 30 autism speech education sessions, interviews with 3 language education experts, and analysis of teaching materials. We developed LingoLift, a GenAI-empowered system, to address barriers to creating coherent lessons tailored to individual learners identified by speech language pathologists. Innovatively, LingoLift leverages the standardized VB-MAPP framework to drive content generation, enabling scientific alignment between children's hierarchical language skills and learning materials. Through an end-to-end architecture, the system bridges lesson preparation and classroom delivery, seamlessly generating and projecting personalized content into the teaching environment.

We then conducted a 3-week deployment study with 10 student groups completing three sessions using LingoLift to evaluate performance and answer RQ2. Results show LingoLift created coherent, effective personalized learning experiences, streamlining complex manual processes, enhancing teaching inspiration, and enabling real-time contextual support. We discovered teachers' unexpected innovative applications and adaptive compensations, for instance, teachers creatively extended the gesture recognition system into immersive games, compensated for digital limitations through multi-sensory aids, and integrated specialized autism intervention strategies for real-time behavioral challenges, highlighting the complementary synergy between educator creativity and AI. We also observed heterogeneous learning experiences among children and discuss inclusive design implications for accommodating diverse learner needs.

In summary, this study contributes:

- (1) A formative study revealing thematic coherence and personalized teaching challenges in autistic children’s oral language teaching practices (presented in Section 3).
- (2) Development of LingoLift, a generative AI-powered system supporting educators’ autism oral language teaching through skill tracking and generation of personalized, coherent lesson plans and visual materials (presented in Section 4).
- (3) A three-week deployment study demonstrating how teachers use LingoLift and how generative AI supports traditional oral language teaching practices (presented in Sections 5,6).
- (4) Design implications for inclusive education contexts, including hybrid physical-digital strategies accommodating diverse learner needs, AI as adaptive collaborative partners enabling real-time classroom responsiveness, and considerations for cultural adaptation and physical deployment (presented in Section 7.2).

## 2 Related Work

### 2.1 Educators’ Personalized Teaching for Autistic Children

Personalized education for autistic children is a comprehensive approach [69, 89], that customizes curriculum content, teaching methods, and learning environments to match each student’s unique linguistic abilities, learning characteristics, and developmental needs [35]. This approach aligns with the highly heterogeneous developmental characteristics of autistic children’s oral language and maximizes their communicative potential [38, 95], which is essential for reducing social isolation and improving life quality [56, 72].

Tailored Individualized Learning Plans are key tools for translating personalized education into practice by defining specific learning objectives, teaching methods and resources [82, 89]. Standardized language assessment tools, including VB-MAPP [94] and ABLLS-R [73], are utilized to **establish personalized objectives**, ensuring achievable and measurable targets that boost student confidence and support continuous learning progress monitoring [82, 99]. Another key strategy involves **incorporating special interests**: autistic children often have strong personalized preferences or intense interests [102]. Including topics and activities that children enjoy increases engagement and motivation [69, 100], leveraging intrinsic motivation to promote natural oral language learning [45]. In summary, effective personalized teaching strategies integrate assessment results, tiered goal-setting, and motivation-driven instruction to create a learning plan that meets individual needs while enhancing learning engagement.

Educators serve an indispensable role in personalized teaching for autistic children [32, 85]. Based on the dynamic assessment of individual children, they develop and continuously refine personalized teaching plans [84], while demonstrating exceptional flexibility in teaching [35] to address the diverse learning challenges encountered by autistic children. Throughout this process, meaningful dialogue and interaction between teachers and children [59, 76] transcend mere knowledge transmission, constituting a vital bridge that facilitates children’s social development [37].

Although personalized teaching significantly improves autistic children’s learning effectiveness, educators encounter substantial implementation challenges, particularly the considerable time investment required for developing and maintaining personalized plans [33, 83] which prevents them from devoting sufficient time to focus on individual needs [81], and the shortage of personalized materials and resources [6, 52]. Therefore, we aim to further explore how to better support educators in implementing personalized autism education while maintaining its personalized quality.

### 2.2 Technology Support for Autistic Children’s Language Development

In recent years, HCI research has demonstrated the potential of digital interactive technologies—including AR [18, 41, 71, 96], VR [11, 49], serious games [13, 43, 44], and mobile applications [62, 101]—in supporting personalized expressive language learning for autistic children. These technologies achieve personalized support through creating engaging training environments [13, 49], facilitating real-world generalization [96, 101], and creating customized material datasets [41, 62].

However, these studies predominantly position children as independent technology users, with limited consideration of educators’ active involvement in the learning process, with some even advocating for technology to replace educators [23, 71]. Although this design enables children to engage in independent repetitive practice [62] and reduces educators’ workload [96], it limits accessibility for children with greater support needs who may struggle to operate digital systems [39]. More critically, as mentioned in Section 2.1, educators play an irreplaceable role in facilitating personalized teaching adaptability and social development [40, 46, 54].

Previous research also tends to focus on singular aspects of language skills, particularly vocabulary learning [18, 25, 41, 62, 96, 101]. According to VB-MAPP, language skills encompass diverse domains including articulation, echoic, mand, tact, linguistic, and intraverbal [94]. Learning to match vocabulary with corresponding referents constitutes only one component of tact skills. While this domain-specific focus provides value for targeted skill development, it limits the comprehensive development across multiple core language domains that children require.

In contrast to previous research, LingoLift emphasizes educator-centered technological support [103] and focuses on comprehensive verbal behavior and oral language development for autistic children.

### 2.3 Generative AI Supports Personalized Teaching Practice

Recent advances in generative AI have enabled new technical support for personalized learning through content adaptation tailored to children’s characteristics [16, 52, 91]. For cognitive-level adaptation, systems like EmoEden and BrickSmart dynamically adjust learning content difficulty and pacing based on children’s developmental stages [51, 55, 97]. For contextual adaptation, OSOS integrates content generation with children’s lived experiences and situational contexts to ensure content relevance [51, 97]. For interest integration, StoryBuddy and EMooly incorporate children’s preferences into generated content to enhance learning engagement [60, 97, 108]. Additionally, these systems can rapidly respond

to children’s behavioral and environmental changes in real-time, generating customized content [14, 55, 107].

Recent research has also emphasized AI’s role in empowering educators [40], particularly in accessing richer educational materials, enhancing performance, and reducing workload [78]. Educators are adopting general-purpose models like ChatGPT and DALL-E to obtain more suitable educational materials [63, 77], addressing challenges of insufficient personalized resources [16, 52]. For busy speech-language pathologists who struggle to dedicate sufficient time to each patient, SPELTA and QuickPic enable rapid development of therapeutic content [24, 81]. LessonPlanner [20] and BrickSmart [55] provide step-by-step guidance for less experienced educators. While these systems demonstrate the potential of generative AI in creating customized educational content, their content generation logic is not grounded in the rigorous assessment-based frameworks essential for autism education [73, 94], and their designs do not target one-on-one oral language teaching scenarios with autistic children [20, 55, 81, 108]. LingoLift addresses these gaps by integrating validated assessment instruments and designing for the complete workflow of personalized autism language teaching.

In summary, Generative AI can rapidly create adaptive educational content that responds to individual learning pace, interests, and developmental needs, addressing the challenges of time investment and resource shortage of educators in personalized oral language teaching for autistic children. These advances inspired the design of LingoLift, an AI-generated system that emphasizes educators’ teaching practice scenarios compared with previous works. LingoLift generates rich oral language learning plans and materials for children with autism who have diverse language abilities and learning interests, enhancing educator efficiency and performance while achieving better learning outcomes.

### 3 Formative Study

To develop AI tools that meaningfully support educators (RQ1), we reached out to a leading local special education school and conducted a formative study examining educators’ current practices, needs, and challenges. We employed three complementary approaches: video analysis of 30 hours’ lessons involving 3 educators and 5 autistic children, teaching material analysis, and expert interviews with 3 speech-language specialists. Participants were recruited through the school’s administrative staff, who distributed recruitment notices to speech-language teachers. All participating children had confirmed autism diagnoses. No participants overlapped between the formative study and field deployment. This study received ethical approval from our institution.

#### 3.1 Data Collection and Analysis

Figure 2 shows our formative study methodology including Video Analysis, Teaching Material Analysis, and expert interviews. Each method contributed complementary and mutually reinforcing insights to our key findings.

**Video Analysis:** We recorded 30 hours of one-on-one oral language teaching sessions with 3 educators and five children aged 6-8 years (all with confirmed autism diagnoses), ensuring minimal disruption to natural classroom dynamics. Two researchers

conducted systematic video coding focusing on: (1) lesson structure and flow, (2) teaching content and pedagogical methods, and (3) children’s responses and individual learning characteristics. To ensure coding reliability, two researchers independently coded 20% of the video corpus (two sessions per child, 6 hours). Cohen’s Kappa coefficient was calculated to assess inter-rater agreement, yielding substantial reliability across three dimensions ( $\kappa = 0.78$ , range: 0.75–0.82). Discrepancies were resolved through discussion, after which one researcher coded the remaining videos.

**Teaching Material Analysis:** We analyzed teaching materials from the institution, which served as reliable and authoritative sources given the school’s substantial student enrollment, extensive teaching experience, and certified professional staff. The materials included VB-MAPP-based assessment systems, skill tracking indicators, exemplary teaching cases, and phonics/vocabulary checklists. These materials formed our development corpus for matching children’s language abilities with appropriate teaching objectives and methods.

**Expert Interviews:** Through video analysis and material analysis, we derived initial insights into personalized teaching content design methods, current teaching challenges, and design concept of context-embedded Thematic Learning. We then conducted semi-structured interviews with three experienced speech-language pathologists, all working in rehabilitation institutions (P1: 28 years old, 9 years of teaching experience; P2: 31 years old, 10 years of teaching experience; P3: 35 years old, 8 years of teaching experience). Each interview lasting approximately 40 minutes, to validate our summarized findings and further understand their perspectives on integrating AI tools into their practice. Two researchers independently conducted open coding on all transcripts in NVivo using inductive thematic analysis [7]. Through iterative discussion, codes were collaboratively refined and organized into themes addressing personalized teaching practices, instructional challenges, and design requirements for AI-assisted tools.

#### 3.2 Formative study results

**3.2.1 Current Teaching Practice Overview.** Based on video analysis and material analysis, we identified the educators’ teaching workflow illustrated in Figure 3. **Pre-teaching:** educators develop lesson plans and prepare materials, primarily flashcards covering articulation, vocabulary, and conversation practice; **During Teaching:** instruction focuses on four core components—articulation training, vocabulary development, grammar exercises, and conversational practice; **Post-teaching:** educators record progress and organize materials for subsequent lessons. This workflow ensures continuous oral language instruction for autistic children.

**3.2.2 Systemic Burden in Lesson Preparation and Management.** AI tools for improving time efficiency emerged as educators’ primary expectation, particularly in lesson preparation and student management tasks.

**Preparing Lesson Plans and Materials.** Educators face substantial time investments in developing individualized lesson plans based on previous learning records and sourcing interest-aligned materials. P3 emphasized the burden of “*spending extensive time collecting materials, printing, and creating PPTs before each class.*” Additionally, they must assess children’s current language abilities

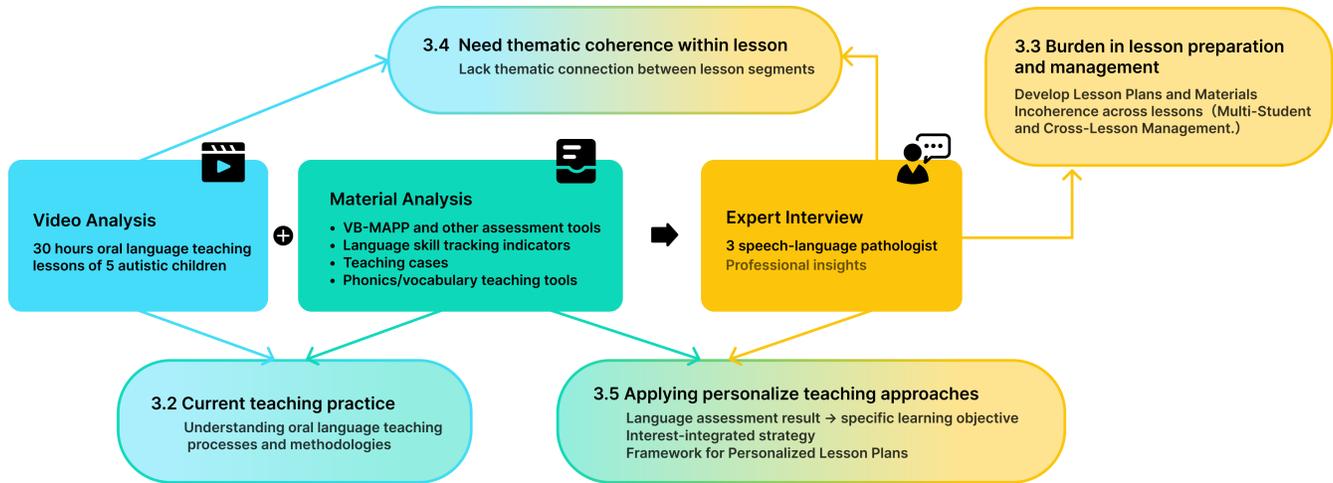


Figure 2: Formative study methodology and key findings.



Figure 3: Workflow and resources in current teaching practice.

and identify knowledge gaps, as P1 noted, “determining what content children have mastered and what they haven’t.” This preparation complexity led educators to express a strong desire to “quickly find appropriate teaching tools and materials to save time” (P2).

**Multi-Student and Cross-Lesson Management.** Beyond individual lessons, educators face management challenges across multiple students and lessons. They organize learning content after class, track individual progress, and conduct ongoing assessments. Maintaining continuity between lessons requires knowledge progression planning, review content selection, and systematic difficulty gradients (P1-P3). P1 stated, “It would be very helpful if AI could automate assistance in tracking children’s learning progress and creating personalized learning pathways.”

**3.2.3 Thematic Coherence Within Lesson.** Video analysis revealed significant learning content incoherence within current teaching practices. As mentioned in Section 3.1, individual lessons comprise multiple skill training components—articulation, vocabulary, grammar, and conversation practice—each using distinct materials and approaches. For example, children might first learn “f” articulation like “airplane” (fēi jī) and “kite” (fēng zheng), then immediately switch to practicing quantity vocabulary like “one schoolbag” and “two dumplings”, with no thematic connection between segments.

This lack of thematic coherence creates abrupt transitions, disrupting learning flow and hindering meaningful concept connections. Interview data confirmed educator’s awareness of this challenge and their desire for cohesive structures. P1 noted that “language intervention could be more systematic if lessons extended around one unified theme,” while P3 emphasized that “integrating activities under a coherent theme would create smoother transitions.” This highlights the potential for AI-assisted tools to develop thematically coherent lessons addressing multiple learning objectives.

**3.2.4 Applying Personalized Teaching Approaches.** Building upon Section 2.1 of teachers’ design and adjustment of personalized teaching content, our formative study further find how these methods are applied in authentic teaching practice.

**Framework for Personalized Lesson Plans.** Based on our analysis of educational materials and classroom observations, teachers’ personalized lesson plan frameworks contain four core components: child profile (language ability and prerequisite skills assessment), personalized teaching objectives, teaching resources, and lesson instructions (detailed step-by-step instructional guidance). This framework “effectively supports teachers in organizing lessons” (P3), providing guidance for our structured lesson plan generation process.

**Ability-Adaptive Teaching Approaches.** Observations revealed significant variations in children’s language abilities—from basic syllables to complex dialogue. Teachers “*identify the essential characteristics of language barriers*” (P1) and employ different teaching content, including: staged phonetic instruction for articulation disorders, vocabulary expansion for limited vocabulary, imitation-to-spontaneous expression transitions, and conversational skill enhancement (P1-P3).

**Interest-Integrated Teaching Strategies.** Teachers employ comprehensive interest-integrated strategies: behavioral reinforcement—“*providing reinforcers as rewards upon goal completion*” (P2, P3), and content adaptation—“*selecting topics of relative interest to children as instructional vehicles*” (P1). For instance, children who enjoyed radish-pulling games were guided through conversations about radish quantity and colors, effectively enhancing learning engagement.

## 4 System Design

To help educators streamline lesson preparation and management workflows, ensure thematic coherence and content integration, and achieve effective personalized teaching, we developed LingoLift, a generative AI-empowered system supporting educators in personalized oral language instruction for autistic children.

Figure 4 illustrates the system workflow: (1) **Child Profile Management**: configuring children’s language abilities and interests through manual initial setup with automatic progress-based updates; (2) **Lesson Preparation**: generating personalized learning plans and materials; and (3) **Lesson Delivery**: seamless integration with projection systems enabling direct use of generated content for teaching.

The system adopts a three-tier architecture (Figure 5): a **Teacher Tablet Client** serving as the educator’s mobile control center; **Backend Services** leveraging RAG and vector databases for intelligent content generation; and a **Classroom Projection Client** enabling gesture-based interaction for end-to-end lesson delivery. Detailed technical implementation is described in Section 4.4.

### 4.1 DF1: Personalized Teaching Content Generation

**4.1.1 Language Ability-Driven Learning Personalization.** LingoLift implements assessment-based personalization aligning individual language capabilities with learning objectives. Educators input child profiles including assessment results across four domains: articulation, tact (naming), linguistic structures, and intraverbal skills. Articulation assessment uses localized Chinese phonetic evaluation protocols based on the “*Articulation and Phonetic Ability Assessment*” framework, while naming, linguistic structure, and conversational abilities use VB-MAPP (Verbal Behavior Milestones Assessment and Placement Program) protocols.

These dimensions correspond to four instructional components: articulation practice, vocabulary acquisition, grammar exercises, and conversational training. Through systematic mapping between language skill domains and constituent tracking sub-items, the system generates stratified learning objectives matching each child’s developmental stage, ensuring appropriate cognitive load while promoting skill advancement across language domains.

**4.1.2 Motivation-Oriented Thematic Personalization.** LingoLift incorporates individual children’s interests and preferred reinforcement stimuli into thematic lesson design. The system maintains detailed interest profiles for each child, enabling generation of contextually relevant learning scenarios. For instance, a child interested in animated content receives lessons themed “*Visiting Doraemon’s Home*,” integrating core activities like “*self-introduction*” and “*identifying household items*” with cartoon imagery and domestic scene materials. This approach transforms abstract learning objectives into concrete experiences leveraging children’s intrinsic motivations.

### 4.2 DF2: Coherent Learning Pathway Generation

**4.2.1 Thematic Coherence Within Lessons.** LingoLift addresses content fragmentation by generating unified material sets maintaining thematic coherence across learning modules. Under a single theme, the system produces materials supporting articulation practice, naming exercises, linguistic structure development, and conversational training, ensuring instructional continuity.

When an educator selects “*supermarket shopping*” with objectives including “g” sound articulation practice, food naming, subject-verb-object sentence training, and polite expressions, LingoLift generates materials serving multiple purposes. A “*cake*” material supports “g” sound practice, enables food vocabulary acquisition, and facilitates role-playing with expressions like “*Hello*,” “*I would like cake*,” and “*Thank you*” within appropriate sentence structures.

**4.2.2 Progressive Learning Across Lessons.** The system maintains learning continuity through automated progress tracking and adaptive content sequencing. Each educator operates through individual accounts managing multiple child profiles. LingoLift records each child’s learning content across sessions, and recommends advancement objectives at increased difficulty levels and review objectives for skill consolidation. This dual-track approach ensures systematic skill building while maintaining mastery of acquired competencies, creating coherent learning trajectories spanning multiple sessions.

### 4.3 DF3: AR Card Interface

LingoLift replaces traditional paper-based materials with projection-based materials addressing content limitations and preparation time constraints. The system substitutes static, commercially-produced cards with dynamically generated digital content, eliminating material availability constraints while reducing educators’ preparation time.

The projection interface operates through synchronized connectivity between projection displays and tablet controllers, enabling content access and manipulation. Educators access personal accounts through the projection system to view and utilize pre-generated lesson plans and materials. Material interaction uses gesture-based controls, allowing real-time content manipulation during instruction without disrupting lesson flow.

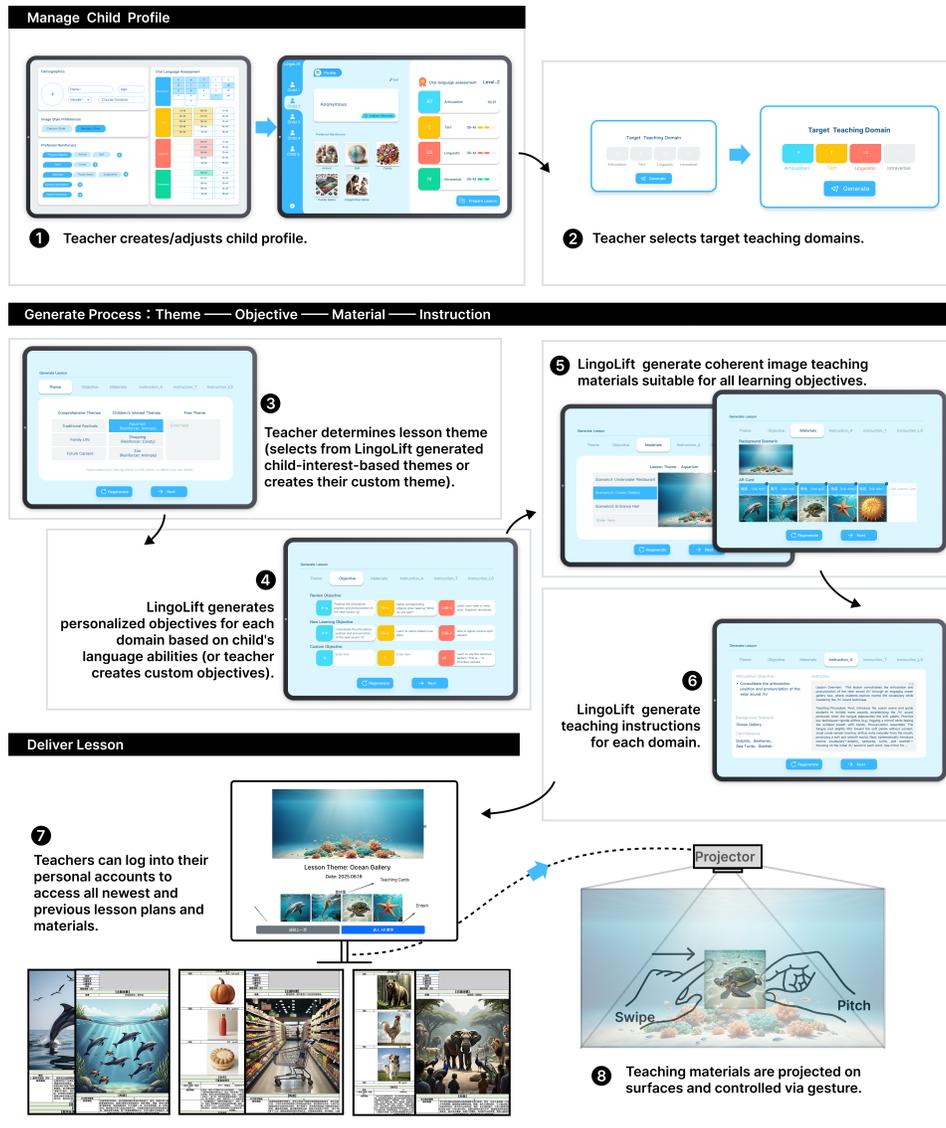


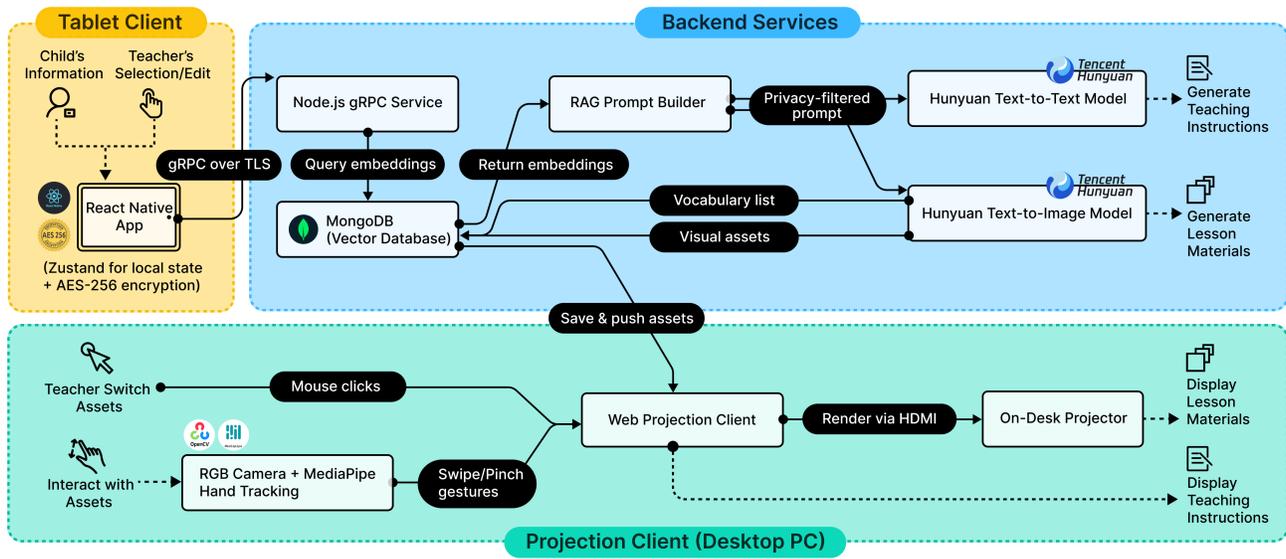
Figure 4: Main interfaces and usage flow of LingoLift.

#### 4.4 System Implementation

LingoLift comprises two coupled clients with a modular back end. The **Teacher Tablet** is a cross-platform **React Native** app where instructors manage child profiles (VB-MAPP level, reinforcement preferences) and select between realistic or cartoon image styles. Local states are stored in **Zustand** and encrypted with **AES-256** [5]; data sync occurs only upon explicit request to preserve privacy.

The **Classroom Projection Client** is a desktop web app linked to the teacher account. Preconfigured lesson plans are rendered as background scenes, with assets controlled via lightweight **MediaPipe** gesture tracking [58] (swipe to move, pinch to lock) or direct clicks, enabling controller-free AR interactions.

On the server side, encrypted **gRPC** [34] requests invoke a **Node.js** pipeline. A **MongoDB** [2] vector store retrieves prior lessons and preferences, which are merged with VB-MAPP goals through a retrieval-augmented generation (**RAG**) pipeline [53]. Few-shot exemplars from expert-written plans are combined with child-specific records into privacy-filtered prompts, then sent to **Tencent Hunyuan** text-to-text [92] and text-to-image models [98]. Generated word lists (with bolded target phonemes) and background illustrations are returned to **MongoDB** [104] and streamed to the classroom client.



**Figure 5: LingoLift System Architecture and Data Flow.** The figure illustrates the three main modules: (1) the Teacher Tablet Client (React Native, Zustand local state, AES-256 encryption), (2) the Backend Services (Node.js gRPC service, MongoDB vector database, RAG prompt builder, Tencent Hunyuan text-to-text and text-to-image models), and (3) the Classroom Projection Client (MediaPipe-based gesture recognition, mouse input, HDMI projection). Data flows are encrypted, privacy-filtered, and tokenized to ensure confidentiality.

## 5 Field deployment

To evaluate LingoLift’s usability, we conducted a three-week field deployment study at a special education school with 10 teacher-student dyads (special education teachers paired with autistic students). The dyads used LingoLift to replace their traditional oral language teaching, completing 30 lessons total (3 per dyad). The study examined how LingoLift influences teachers’ lesson preparation and teaching processes, as well as students’ engagement and learning outcomes through three consecutive oral language lessons per dyad. All participants provided informed consent, with children’s consent obtained through their parents. This study was approved by the institutional review board.

### 5.1 Participants

We conducted school-wide screening to identify autistic children with oral language learning needs and their assigned teachers, recruiting 10 teacher-student dyads. All teachers were the children’s regular instructors, ensuring established familiarity and rapport within each dyad. Child participants (C1–C10; Table 1) were diagnosed with autism spectrum disorder and presented with oral language developmental delays. Recruitment criteria required emotional stability, ability to understand basic instructions, and good cooperation.

Children’s language ability levels were assessed by their assigned teachers based on VB-MAPP, spanning Level 1 (basic vocabulary and simple requests,  $n = 3$ ), Level 2 (short phrases and expanded comprehension,  $n = 4$ ), and Level 3 (sentence formation and conversational skills,  $n = 3$ ). For children who had not yet developed linguistic and intraverbal skills, we marked their abilities as N/A.

Teacher participants (T1–T10) (Table 2) were special education teachers or rehabilitation therapists with professional teaching experience ( $M = 4.9$  years). We also collected self-reported typical lesson preparation time for each teacher.

### 5.2 Study Setup

The study was conducted in a dedicated language training classroom at the special education school. The school has multiple speech training rooms, with each child typically attending lessons in their familiar assigned classroom and teachers rotating between rooms to instruct different students. For this study, we renovated one classroom and equipped it with fixed experimental equipment, including network connectivity, projection facilities, cameras, and gesture recognition devices (see Figure 6). All participating dyads conducted their lessons in this renovated room on a scheduled basis. Due to equipment constraints, we adopted a one-teacher-one-student design where each teacher worked with only one child in a fixed classroom throughout the study, minimizing environmental disruptions for the children.

Each participating teacher received a Huawei tablet pre-installed with the LingoLift application and configured with individual user accounts. Teachers could flexibly use the tablets for lesson preparation, whether at home or in their offices.

### 5.3 Procedure

The study was a three-week field deployment divided into three phases: onboarding workshop, field deployment, and evaluation.

**Phase 1: Onboarding Workshop.** We conducted an on-site workshop at the special education school, introducing the study

**Table 1: Child participant demographics.**

ID	Age	Gender	Articulation	Tact	Linguistic	Intraverbal	Reinforcers
C1	8	Female	b, p, m, g, j, q, r, z, f, l	9-M (Lv2)	8-M (Lv2)	8-M (Lv2)	Toy phone; lollipop; sensory room; nursery music
C2	9	Male	b, p, m, f, d, t, l, j, h	12-M (Lv3)	13-M (Lv3)	8-M (Lv3)	Toy cars; candy; jelly; milk; fruit
C3	7	Male	b, p, m, z, j, d, t, zh, r, g, f, h	8-M (Lv2)	8-M (Lv2)	6-M (Lv2)	Puzzles; milk; drawing; mirror; role-play; cookies; wooden cards; candy; Lego; rhymes; iPad
C4	7	Male	b, p, m, d, h, g, k, n	5-M (Lv1)	6-M (Lv2)	6-M (Lv2)	Spinning toys; fruit; bubbles; tickling; high-fives
C5	8	Male	b, p, g, zh, d, n, x, r, k, m, t, h, q, j	2-M (Lv1)	N/A	N/A	Shrimp crackers; corn; Peppa Pig; sensory toys; ball games
C6	8	Male	m, x, zh, g, k, s, z, d	3-M (Lv1)	N/A	N/A	Toys; chips; free games
C7	6	Male	b, m, x, d, j, g, h, k, q	8-M (Lv2)	7-M (Lv2)	6-M (Lv2)	Ball; shrimp crackers; tickling; visual stimuli; Peppa Pig
C8	9	Female	b, p, m, j, x, g, h, t, n, z	4-M (Lv1)	6-M (Lv2)	6-M (Lv2)	Peanuts; tactile stimuli; eye contact; leaves; twigs; dates; raisins
C9	8	Female	b, m, j	2-M (Lv1)	6-M (Lv2)	6-M (Lv2)	Lollipops; hawthorn sticks; shrimp crackers; BabyBus stories; watching TV
C10	8	Female	b, m, p	1-M (Lv1)	N/A	N/A	Cookies; arranging identical items

**Table 2: Teacher participant demographics.**

ID	Age	Profession	Teaching Experience	Students Instructed
T1	27	Special Education Teacher	4 years	30
T2	27	Special Education Teacher	4.5 years (0.5 at this school + 4 at hospital)	2
T3	32	Special Education Teacher	5.5 years (3 at this school + 2.5 at hospital)	10
T4	24	Special Education Teacher	2 years	8
T5	27	Special Education Teacher	5 years (3 at this school + 2 at hospital)	20
T6	28	Special Education Teacher	7.5 years (3 at this school + 4.5 at rehabilitation center)	12
T7	24	Rehabilitation Therapist	3 years	50+
T8	26	Rehabilitation Therapist	3.5 years	25
T9	39	Special Education Teacher	8 years (3 at this school + 5 at hospital)	8
T10	29	Special Education Teacher	6 years (1 at this school + 5 at another school)	2

procedure and providing LingoLift usage training. Each teacher received a dedicated tablet and created their user account. Teachers practiced preparing a sample lesson plan to ensure proficiency with LingoLift. We also provided instructional videos for reference.

**Phase 2: Field Deployment.** During the three-week period, teachers used LingoLift for lesson preparation and teaching. Each teacher-student dyad completed three oral language lessons following their curriculum schedule (30 lessons in total). To avoid disrupting the natural teaching environment, we did not enter classrooms but recorded lessons using mounted cameras and provided support online when needed. Teachers completed a post-lesson questionnaire (see Appendix A) after each session to reflect on their teaching experience and system usage.

**Phase 3: Evaluation.** Teachers filled out a post-study questionnaire (see Appendix B) after completing all lessons, assessing usefulness, usability, and satisfaction with LingoLift. We then conducted 40-minute semi-structured interviews with each teacher. The interviews covered four themes: (1) detailed discussion of their post-lesson questionnaire responses, (2) overall experiences with using LingoLift for lesson preparation, (3) evaluation of LingoLift’s personalization features and their effectiveness, and (4) reflections on the teaching process when incorporating LingoLift (see Appendix E).

## 5.4 Data Collection and Analysis

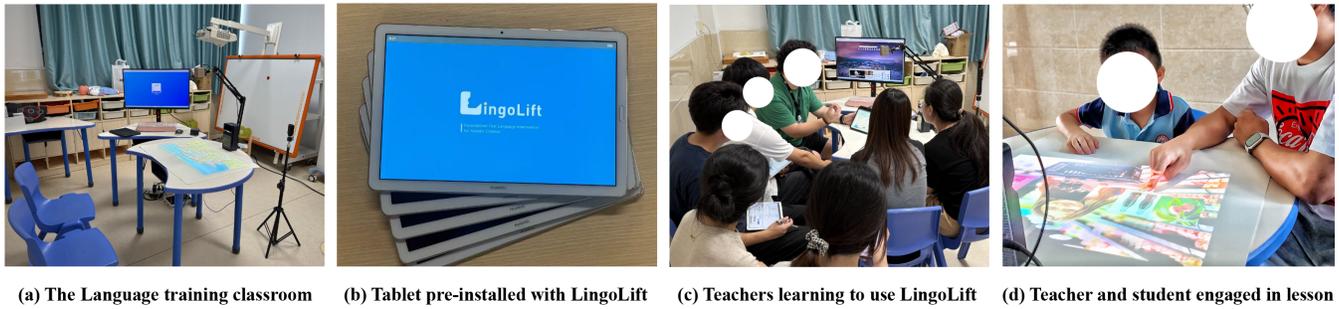
We employed a mixed-methods approach combining quantitative and qualitative analyses. Our analysis incorporated three primary data sources:

**5.4.1 Self-reported Questionnaires.** We collected feedback through a post-lesson questionnaire after each lesson (30 total responses) and a post-study questionnaire at the study’s conclusion (10 total responses).

The **post-lesson questionnaire** was informed by CIPP evaluation [90] and AI-generated content perception [87] frameworks. It comprised 16 items using 7-point Likert scales to measure: (1) Personalized Alignment—how well the lesson matched individual student needs; (2) Lesson Coherence—the logical flow within individual lessons and continuity across sequential lessons; (3) AI Satisfaction—teachers’ satisfaction with AI-generated content; and LingoLift-specific dimensions: (4) Engagement and (5) Learning Outcomes. The questionnaire demonstrated good reliability (Cronbach’s  $\alpha = 0.820$ ).

The **post-study questionnaire** integrated elements from the Mobile-health App Usability Questionnaire (MAUQ [109]) and Technology Acceptance Model (TAM [15]), employing 17 items with 7-point Likert scales to evaluate LingoLift system performance, specifically examining perceived usefulness, ease of use, and user satisfaction.

**Data Analysis.** Statistical analyses included non-parametric tests and visualization of distributions. Teachers’ questionnaire responses (Q1–Q17) were aggregated into four composite dimensions (Usefulness–Efficiency, Usefulness–Performance, Ease of Use, Satisfaction). For each teacher, we computed mean scores per dimension and derived descriptive statistics (mean, SD, median, min–max). Scores were then transformed into long format to facilitate dimension-wise comparisons and distributional visualization.



**Figure 6: Photos from the field deployment study. (a) The language training classroom used for the study. (b) Tablets pre-installed with LingoLift. (c) Teachers learning to use LingoLift. (d) Teachers and children engaged in lesson activities.**

Weekly evaluations (Q1–Q16) were analyzed by aggregating items into pedagogical dimensions and applying repeated-measures tests across Weeks 1–3.

**5.4.2 Interview Transcripts.** We audio-recorded all 10 semi-structured interviews and transcribed them verbatim with subsequent content verification. To analyze the qualitative data, two researchers adopted a **reflexive thematic analysis** [8], which emphasizes iterative engagement and researcher reflexivity. Both researchers independently conducted open coding on three representative transcripts, noting initial descriptive codes. Through iterative discussion and memo-writing, they collaboratively refined these into a unified coding framework. This process yielded a three-level coding structure: (1) open codes capturing salient concepts in participants’ words, (2) axial codes grouping related concepts into broader categories, and (3) selective themes synthesizing these categories into higher-level interpretations relevant to the research questions. After establishing the shared coding scheme, the researchers applied it across all transcripts in **NVivo**, while remaining open to inductive refinements. Regular peer-debriefing sessions were held to negotiate differences, reflect on positionality, and ensure interpretive rigor. This multi-stage procedure enhanced both coding consistency and the trustworthiness of the analysis.

**5.4.3 Video Observation and System-Generated Content Analysis.** We recorded 30 lessons and conducted video observation, supplemented by analysis of LingoLift’s generated teaching content (learning objectives, lesson materials, and teaching plans). Video observations examined children’s learning outcomes, engagement, and teachers’ teaching methods to assess LingoLift’s effectiveness in content generation and teaching guidance. To better understand how teachers worked with the system to obtain these materials, we also exported logs of their prompts for both text generation (e.g., stories, questions) and text-to-image generation during the deployment. Two researchers engaged in a reflexive thematic analysis [8] of these prompt histories and their surrounding notes, iteratively discussing and refining interpretive codes that captured how teachers adjusted prompts to express pedagogical intentions, local contexts, and children’s sensory needs. The integration of video observations, prompt histories, and generated content analysis enabled triangulation of findings.

## 6 Findings

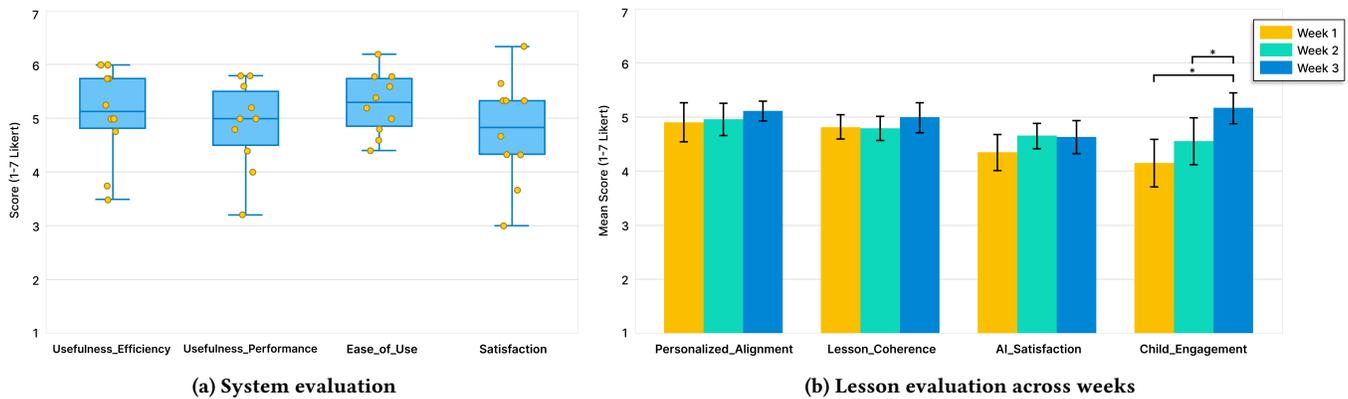
Our field deployment demonstrated that LingoLift exhibits good usability in educators’ practices, supporting children’s personalized learning and language skill development. AI content generation enhanced traditional teaching practices by simplifying complex manual processes and enhancing teaching inspiration, enriching learning materials and enabling flexible content adaptation. Beyond its intended functionality, teachers innovatively extended LingoLift’s capabilities through creative adaptations and multi-sensory compensations.

### 6.1 System Evaluation

As shown in Fig. 7a, teachers’ ratings of LingoLift were consistently above the neutral midpoint of 4 on a 1–7 Likert scale. **Ease of Use** received the strongest consensus ( $M = 5.28$ ,  $SD = 0.58$ , median= 5.30), with scores clustering between 5 and 6, indicating successful integration of AI features into intuitive interfaces with low adoption barriers. Both dimensions of **Usefulness**—efficiency ( $M = 5.08$ ,  $SD = 0.88$ ) and performance ( $M = 4.88$ ,  $SD = 0.83$ )—centered around 5. The system reduced preparation time from  $M = 72$  minutes (range: 30–150) to  $M = 24.7$  minutes (range: 10–60) (66% reduction) while maintaining high perceived work quality. **Satisfaction** exhibited greater variability ( $M = 4.80$ ,  $SD = 1.00$ , min= 3.0, max= 6.3). While the median remained favorable at 5.0, several low outliers revealed a polarized pattern, reflecting interview accounts where teachers reported occasional friction when AI-generated images did not fully align with their pedagogical intentions. In particular, some teachers described that obtaining highly situated or locally recognizable scenes could require additional prompt refinements and manual selection, which tempered satisfaction for a subset of participants despite overall positive views. Taken together, these results highlight LingoLift’s stable baseline of usability and efficiency.

### 6.2 Lesson Evaluation

**6.2.1 Quantitative Result.** As shown in Fig. 7b, lesson evaluation scores remained consistently above the scale midpoint (4 on a 1–7 Likert scale), indicating that the system reliably delivered pedagogically aligned and coherent lesson content. While a slight upward trend was observed across weeks for most measures, no



**Figure 7: System evaluation and weekly lesson evaluation. (a) Box plots of post-study questionnaire results (1-7 Likert); individual data points and box plots display response distribution with mean  $\pm$  SD. (b) Teacher post-lesson ratings for four dimensions across Week 1-3 with error bars indicating standard error. Asterisks (\*) denote statistically significant differences at  $p < 0.05$ .**

statistically significant week-to-week changes were detected for *Personalized Alignment*, *Lesson Coherence*, or *AI Satisfaction*.

### 6.2.2 Qualitative Result.

*Personalized Matching of Learning Objectives.* Based on VB-MAPP assessment results, LingoLift successfully created learning objectives that matched individual children’s needs and abilities. As T10 noted in the interview: “*The content it generates aligns well with students’ individual learning traits and needs,*” ensuring “*learning objectives matched children’s abilities*” (T1-4, T9). Moreover, LingoLift demonstrated dynamic adaptation over time, with educators observing “*continuous advancement in learning objective difficulty across sessions*” (T1) while maintaining pedagogical coherence and progression.

*Rich and Adaptive Teaching Materials.* Another key strength of LingoLift was “*generating diverse, digitally-enhanced topics and materials tailored to each child’s personalized interests and needs*” (T1, T3-T5, T8), overcoming the limitations of traditional teaching resources (T3). Teachers (T4-6, T8, T10) highlighted how the system eliminated financial and logistical constraints associated with physical materials. As T10 explained: “*Previously, preparing physical objects lacked flexibility due to cost and effort... but with digital images, I can create abundant materials and prepare multiple rich lessons whenever needed.*”

## 6.3 Children’s Learning Performance

**6.3.1 Learning Engagement.** As shown in Figure 7b, children’s engagement demonstrated significant improvement over three weeks (Week 1:  $M = 4.20$ ; Week 3:  $M = 5.10$ ,  $p < 0.05$ ), indicating progressive adaptation to LingoLift. The sustained engagement growth suggests that LingoLift maintained student interest over the study period, rather than showing the decline often associated with novelty effects.

**6.3.2 Learning Outcome.** Teachers evaluate each lesson on a 1–7 scale to assess whether learning objectives were achieved and

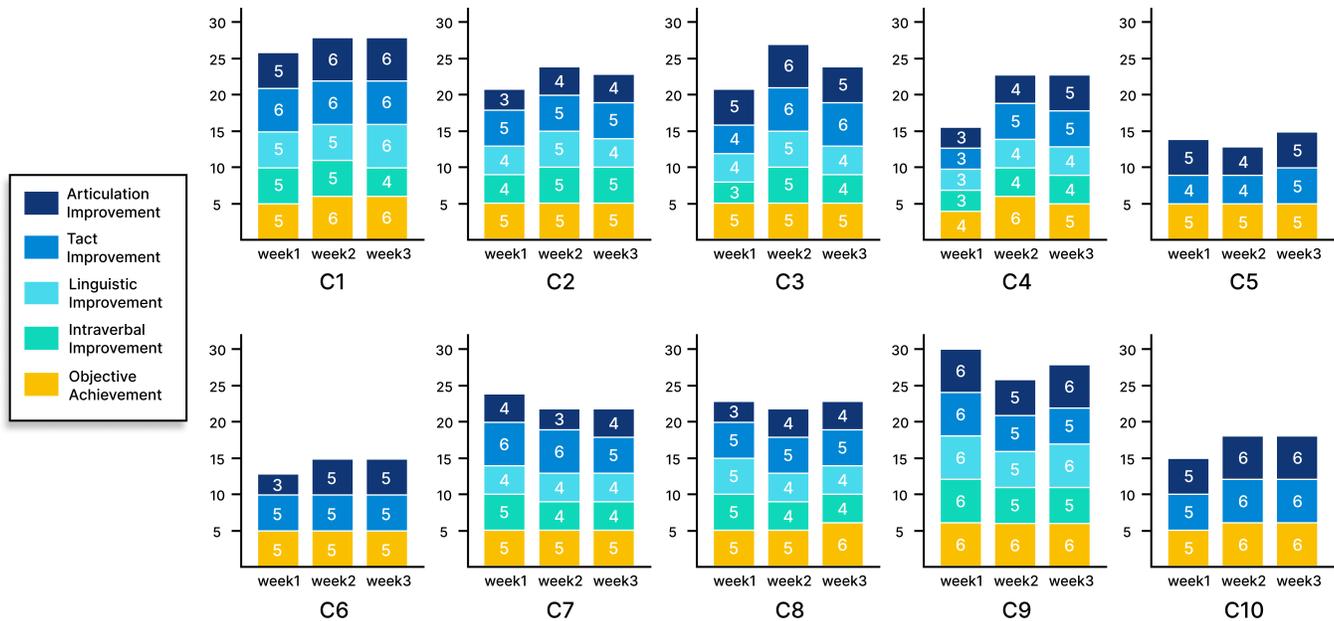
the degree of improvement in five language skill categories (see Appendix A). As shown in Figure 8, children successfully achieved the established classroom learning objectives ( $M = 5.27$ ,  $SD = 0.52$ ), indicating high agreement that learning objectives were met, demonstrating that LingoLift generated appropriately personalized learning content.

Teachers rated the improvement in tact skills most highly ( $M = 5.17$ ,  $SD = 1.03$ ), reflecting LingoLift’s advantages in providing visual-contextual learning support for concrete language concepts. Teacher-rated improvement in linguistic and intraverbal remained minimal, with teachers explaining that “*advanced language skills are difficult to produce quantifiable progress within 3 weeks*” (T5, T7).

### 6.3.3 Individual Differences in Children’s Learning Experience.

While aggregate engagement scores improved over time, video observations and teacher interviews revealed substantial individual differences in how children experienced and interacted with the system. Engagement patterns showed polarization among participants, reflecting individual differences in learning preferences and cognitive processing styles.

**High Engagement and Positive Adaptation.** Seven children (C1-C4, C7-C9) demonstrated strong positive responses to LingoLift throughout the study. Video observations captured these children leaning toward projections, spontaneously reaching to interact with virtual objects, tracking visual materials with their gaze, responding promptly to teacher instructions, and maintaining focus for extended periods (15–20 minutes per session). These high-engagement children typically exhibited three characteristics: (1) **Digital interface adaptation**, demonstrated by their ability to recognize digital materials and understand interaction rules—for instance, C1 spontaneously repositioned projected images to appropriate locations in the third lesson, explaining “the rooster is on the tree”; (2) **Virtual scenario engagement capacity**, particularly for children with imaginative play and role-playing abilities, enabling deep immersion in system-generated contexts and rich interactions with teachers, such as C3 engaging in simulated shopping activities (“*Let’s put Fanta into our basket... Now we’re going to checkout*”); (3)



**Figure 8: Teacher-rated learning outcomes over three weeks. Left panel: teacher ratings of improvement (1–7 scale) for individual children, where 1 = No improvement, 4 = Moderate improvement, and 7 = Significant improvement.**

**Intrinsic motivation triggered by personalized content.** As T3 observed: “My student was particularly interested in animals, so with the zoo-themed materials the system generated, he could focus much better on classroom activities.” T8 also noted: “When the system showed his favorite cartoon characters, he would immediately engage and try to complete all tasks.”

**Engagement Challenges and Barriers.** Conversely, some children exhibited persistent engagement difficulties. Video observations showed these children struggling to focus on digital content, frequently looking around, fidgeting with the projector and tablecloth, showing no response to teacher instructions, or producing emotional vocalizations. Teacher interviews identified three primary sources of these challenges: **(1) Difficulty perceiving virtual content.** T6 explained: “Understanding virtual images is challenging for some autistic children.” In one session, C10 repeatedly touched the wall behind projected images, seemingly unable to attend to the virtual layer (T10). T3 and T4 attributed this to cognitive readiness, noting “electronic products suit children with better cognitive abilities” and “they need to understand digital interaction rules first.” **(2) Preference for physical manipulatives.** These children showed stronger engagement with physical objects than virtual representations. For example, T5 interspersed bubble-blowing toys when teaching bubble vocabulary, significantly improving C5’s engagement. However, physical materials did not universally solve engagement issues. When T6 switched from picture-based to physical toy-based teaching for C6, the child became overly absorbed in playing with toys, ignoring all instructional content and showing severe emotional resistance when the teacher attempted to redirect attention—a worse outcome than before. T3 summarized: “The system works well for most children, but for some, we need options

to blend digital and physical materials.” **(3) Initial adaptation challenges.** Equipment appearance had unexpected impacts on children’s responses. C5 initially showed fear of entering the classroom, with T5 noting: “The camera stand looked somewhat like a spider, which he’s afraid of.” C4 and C8 showed curiosity toward the projector, becoming distracted or exhibiting destructive behaviors during the first lesson. These issues notably improved by the second and third lessons, indicating increased adaptation over time.

## 6.4 LingoLift Enhancing Traditional Teaching Practices

LingoLift enhanced traditional teaching practices by streamlining complex manual processes, enhancing content inspiration, and enabling real-time contextual support.

**6.4.1 Streamlining Complex Manual Processes.** Traditional autism education preparation required teachers to navigate multiple platforms and documents, then manually integrate scattered resources—a labor-intensive process. LingoLift systematically integrates this workflow: “Now I can handle all preparation work with just a tablet” (T2). All teachers (T1-T10) reported significant time savings. T2 noted: “Previously, content creation was extremely time-consuming, and I could barely prepare systematic lesson plans for each class, but now I can.” Multiple teachers (T2, T5, T7, T8) emphasized how the system replaced time-intensive resource gathering, with T5 reflecting: “I used to search for teaching resources myself... which required substantial time investment.” T3 added that “digital material generation eliminated labor-intensive manual tasks.”

**6.4.2 Enhancing Teaching Inspiration.** Traditional teaching constrained teachers’ content creation within personal knowledge

boundaries and creative thinking. Teachers agreed LingoLift leveraged computational intelligence to provide enhanced content inspiration (T1, T3, T5, T7, T8, T9). T3 explained: “*The generated vocabulary benefits from AI beyond human cognitive capacity, producing content I wouldn’t have conceived.*” This support not only enriched teaching materials but also transformed teachers’ creative processes. T8 reflected: “*I can conveniently select and adjust from generated content, which is much simpler than starting from scratch.*” T1 directly stated: “*I rely on the content generated by this system.*”

**6.4.3 Enabling Real-Time Contextual Support.** Traditional teaching practice was constrained by pre-made materials, unable to flexibly adjust during lessons. LingoLift addressed this limitation through real-time content generation. T3 shared: “*When I think of new vocabulary or notice student changes, I can quickly generate needed content without being restricted by pre-made materials.*” Additionally, teachers (T1, T3, T7, T10) recognized LingoLift’s thematically coherent materials “*created situational support for children’s learning,*” transforming teaching practice from material presentation to dynamic context creation, supporting skill transfer into the real world (T1, T10).

## 6.5 Teachers’ Adaptive Teaching Innovation

Teachers demonstrated adaptive teaching innovation with LingoLift, going far beyond basic system usage. Their adaptive strategies encompassed creative functional extensions, multi-sensory compensations, and comprehensive integration of specialized autism intervention techniques.

**6.5.1 Creative Extensions Based on System Functions.** LingoLift’s gesture recognition technology was initially designed for intuitive control of learning materials, but we observed teachers’ creative extensions that significantly enriched the learning experience. For instance, T2 created an immersive “*parking lot game*” utilizing parking lot background images and vehicle card images, which promotes spatial reasoning abilities alongside language acquisition. T1 adapted the gesture recognition to design a “*butterfly chasing*” activity, where students’ hand movements became engaging classroom interactions, embedding language learning within motor skills training and imaginative play. These innovations represent more than technological adaptations—they constitute pedagogical creativity that transcends the system’s predetermined functional boundaries.

**6.5.2 Multi-sensory Teaching Compensation for System Limitations.** While LingoLift provided rich visual learning materials, teachers identified “*inherent limitations of digital systems*” (T6) and proactively prepared multi-sensory teaching aids to complement the digital content. For example, teachers addressed auditory limitations by incorporating video materials with authentic animal sounds during zoo-themed lessons (T1), provided tactile experiences through water basins for exploration during ocean-themed sessions (T7), and enhanced olfactory and gustatory learning via real objects during food naming activities (T5). Additionally, T9 employed mirrors and pinwheels to facilitate observation of articulatory movements and airflow during pronunciation practice. T5 articulated the theoretical rationale: “*For children with cognitive difficulties, sensory experience must precede abstract learning. Neural reflexes established through*

*direct perception enable more effective transition to card-based materials.*”

**6.5.3 Negotiating AI-Generated Content via Prompt Refinement.** Our analysis of teachers’ prompts revealed that educators actively negotiated AI outputs rather than accepting them as-is, iteratively refining prompts to align content with pedagogical intentions, local contexts, and children’s sensory needs. For instance, T2 wanted to rehearse public transport routines using pictures of her city’s distinctive buses so that classroom practice would carry over to everyday commuting; she repeatedly specified the city name, bus type, and nearby landmarks, adjusting the wording more than ten times, yet early generations still showed generic or foreign-looking vehicles and streetscapes, so she eventually “settled for the closest one” and supplemented it with school-provided photos. A similar negotiation unfolded around what counted as visually and emotionally appropriate content: when preparing a history museum lesson, T7 found that the model occasionally introduced unsettling elements such as mummies and dark, horror-like exhibits that she felt were unsuitable for autistic children who were easily frightened. In response, she rewrote prompts to explicitly request “bright, friendly museum halls without scary objects,” and manually filtered out any remaining images that might overwhelm children. Across cases, teachers described toning down overly colorful, cluttered, or stylized images by asking for “simple, clean backgrounds” or “calm faces,” and curating only those outputs that felt safe and pedagogically meaningful. These practices illustrate how teachers appropriated prompt design as an interactional layer on top of LingoLift, actively steering AI-generated content toward contextually relevant and emotionally appropriate materials.

**6.5.4 Comprehensive Application of Professional Strategies.** Video observations and interviews revealed that teachers employed language teaching strategies in their instructional practices that aligned with LingoLift’s generated pedagogical guidance: (1) **Modeling and Imitation**, where teachers provided clear articulatory demonstrations for children to observe and imitate speech movements; (2) **Prompting and Responding**, where teachers displayed picture assets and used structured questioning (e.g., “What is this?” “What is he doing?”) to guide language production; (3) **Sentence Completion**, articulating partial sentences for children to supply key vocabulary (e.g., T7: “In the ocean there are...” prompting the response “sea turtles”); and (4) **Instruction Following and Spatial Language Training**, combining gesture recognition with directional commands (e.g., “Put the dog under the tree”) to reinforce spatial vocabulary and listener responding skills.

Beyond these foundational strategies, teachers demonstrated exceptional pedagogical flexibility, integrating comprehensive autism behavioral intervention strategies to address real-time behavioral and learning challenges (T4). For attention deficits, teachers employed pivotal response training and high-frequency reinforcement. T10 noted: “*My student requires very high-frequency reinforcement to maintain focus on learning.*” T5 used exaggerated vocal tones and visual stimulation to capture attention and redirect focus. For emotional regulation, T3 implemented systematic protocols including spatial repositioning and system adjustments: “*When children become overly excited, I move them away from the operational area and turn off the projection system if necessary.*” Teachers

also emphasized flexible lesson adaptation. T6 explained: “*The lesson plan serves as reference, but implementation depends on classroom performance, requiring real-time adjustments to instructional language and sequence.*”

## 7 Discussion

### 7.1 Integrating Generative AI for Autistic Children’s Oral Language Learning

Existing research has explored personalized applications of generative artificial intelligence in emotional learning [60, 97], social skill development [3, 36], and life skill training [9] for autistic children. Recent advances have also demonstrated AI’s potential to empower educators through tailored content generation [20, 55, 81, 108]. However, our research reveals that oral language learning for autistic children possesses unique characteristics that require specialized design considerations beyond generic educational content generation.

**Generation Based on Standardized Assessment.** LingoLift adopts an assessment-driven generation architecture grounded in clinically validated language assessment frameworks—VB-MAPP [94] and Chinese Articulation Test. This approach differs from systems not specifically designed for special education, such as LessonPlanner [20], which generates content based on Gagne’s nine instructional events, or StoryBuddy [108] and BrickSmart [55], which transform children’s natural language descriptions into generated content. This assessment-driven approach enables **Hierarchical skill decomposition**, mapping assessment results to fine-grained learning objectives that align with clinical frameworks, and **Progress-adaptive sequencing**, automatically recommending progressive and review objectives across lessons based on individual mastery levels. This architecture ensures generated content closely matches the assessment-based learning needs characteristic of autism education [83, 84, 99].

**Hierarchical Progression and Logical Matching Requirements.** Speech skill learning objectives are typically specific and quantifiable, exhibiting interconnected hierarchical progression characteristics [48]—from pronunciation and vocabulary to grammatical structures and conversational abilities—where each skill serves as a component of the next [47], forming a nested structure like boxes within boxes rather than parallel independent goals. Therefore, content generation must consider the progressive logical relationships or conflicts between skills [68]. For example, learning color naming of objects aligns well with the grammatical structure of adjectives modifying nouns, but conflicts logically with learning tense expressions (“*I have done*”/“*I will do*”), making effective integration [66] within the same lesson challenging. Thus, even when recommending from learning objectives at the same level, particular attention must be paid to logical relationship compatibility verification [80].

**Situated Learning Context Generation.** Language learning for autistic children requires rich contextual support to facilitate understanding and application [50, 76, 105]. However, unlike emotional learning, social skills, and life skill training—which inherently possess contextual situational attributes—language training is more complex in constructing scenarios that satisfy multiple learning objectives. For example, when teachers select articulation learning

objectives (learning the /g/ sound), vocabulary (food naming), grammar (subject-verb-object patterns), and conversation skills (learning polite expressions), they need to integrate these different dimensional skills into a single coherent learning experience while additionally considering children’s motivational interests. Under numerous requirement constraints, generative artificial intelligence’s computational power beyond human cognition makes it particularly suitable for addressing this integration challenge [12, 29]. Building upon content generation, LingoLift’s end-to-end architectural design bridges classroom preparation and deployment, further supporting such contextualized learning experiences.

Addressing these unique characteristics of language learning, LingoLift employs a step-by-step framework of “*generate learning theme* → *target domains* → *generate learning objectives* → *generate visual materials* → *generate instructional guidance*” [20], progressively refining generated information to ensure accuracy and rationality of the generation process while allowing teachers to make precise adjustments at each stage. This framework achieves situated thematic learning, effectively integrating multi-dimensional language objectives into coherent learning situations.

### 7.2 Design Implications for Inclusive Education Contexts

**7.2.1 Designing Hybrid Physical-Digital Support Strategies.** Our findings reveal heterogeneity in children’s engagement (as shown in Section 6.3.3), some children performed well with LingoLift, others required additional sensory experiences to establish foundational learning, confirming prior work [42, 110] that emphasizes the necessity of physical-digital synergy in special education. In response, we suggest that future special education system designs could benefit from implementing **digital readiness assessment and progressive adaptation**, including evaluating children’s digital readiness across their understanding of digital interaction conventions, imaginative engagement with virtual scenarios, and adaptation speed to novel technological stimuli. Systems could adapt content delivery based on each child’s capability—identifying which materials suit purely digital presentation and which may benefit from physical supplementation (e.g., abstract concepts, sensory-rich vocabulary), and facilitating gradual transitions toward digitalized presentation as children demonstrate increased adaptability and comprehension. Additionally, **multimodal sensory congruence** could ensure that for content flagged as requiring additional support, systems recommend complementary tactile, olfactory, or auditory materials to teachers, maintaining accessibility for children with diverse sensory learning preferences and supporting educators in orchestrating coherent cross-modal learning experiences.

#### 7.2.2 Defining AI as Adaptive Collaborative Partners.

**Real-Time Adaptive Capabilities.** Prior research has established educators’ indispensable role in dynamically adapting personalized teaching for autistic children [32, 35, 84, 85]. Our findings in Section 6.5.4 provide empirical evidence for how AI systems can augment this adaptive capacity in unpredictable special education contexts, such as dynamically generating new vocabulary and adjusting projection interactions based on emotional states. This suggests defining AI as an adaptive collaborative partner that

responds to the unpredictable, high-variability nature of special education classrooms. We suggest three key capabilities to support this role: **generative flexibility for hot-fixes** that enables educators to maintain instructional fluency through rapid on-demand content generation (as demonstrated by LingoLift); **behavioral co-regulation** that actively participates in classroom climate management by adjusting visual complexity, stabilizing interactions, or introducing dynamic audiovisual cues to address hyperactivity or recapture attention; and **embedded expert guidance** that integrates domain-specific special education knowledge to provide real-time strategic support for educators during complex behavioral episodes.

**Facilitating Long-term Reciprocal Co-evolution.** Beyond immediate adaptation, our observations highlight the potential for a long-term reciprocal co-evolution between AI and teachers. We suggest that future systems could be designed to support a functional complementarity where human expertise and machine capabilities mutually reinforce one another: **AI as a Scaffold for Cognitive Externalization:** AI systems can help translate teachers' tacit professional judgment [26] into actionable, structured content. This approach supports teachers by reducing the cognitive load of material preparation. **Leveraging Contextual Translation for System Optimization:** Conversely, as teachers perform "contextual translation" by reorganizing AI outputs based on real-time needs, these interactions offer rich insights for system optimization. To foster co-evolution [21, 22, 31], future designs might move beyond one-way generation to implicitly learn from these creative adaptations. By interpreting teacher behaviors—such as prompting strategies and improvisational adjustments—as feedback signals, the system could refine its generation quality over time, promoting a virtuous cycle of continuous improvement for both the human and AI.

**7.2.3 Localized and Culturally Adaptive AI-Generated Content.** The generative capabilities of AI systems present both opportunities and risks for cultural representation in special education. While generative AI offers unprecedented scalability, it risks perpetuating dominant cultural narratives [4], creating barriers to real-world transfer for children from diverse backgrounds [67], particularly for autistic children who frequently experience difficulties with generalization [10, 64]. Building on the teacher adaptations observed in Section 6.5.3, we propose a design recommendation for **culturally responsive generation**—future systems should move beyond mass production to thoughtfully integrate local knowledge, community values, and individual contexts. By creating materials that resonate with children's lived experiences and cultural realities, AI-generated content can serve as a bridge to meaningful engagement and real-world application.

**7.2.4 Physical Deployment Considerations.** Based on our observations of children's adverse reactions, maladaptive behaviors, and teacher control strategies, we recommend that researchers and designers **conduct sensory audits of physical equipment**, evaluating visual, auditory, and tactile properties that may trigger fear or distraction in sensory-sensitive populations. Thoughtful physical design can significantly enhance children's adaptability

and engagement [27, 106]. **Design for spatial flexibility**, providing configurable hardware options (e.g., adjustable projection locations, mobile versus fixed installations) to modulate student proximity while accommodating teachers' needs for physical access to guide attention and behavior. As prior research indicates, appropriate spatial arrangement is a critical factor in successful deployment [75, 86].

### 7.3 Limitation and Future work

**Deployment Context.** Our deployment involved experienced teachers in well-resourced cities, yet support levels for children with autism vary significantly across different regions. Future research should examine system performance across different socioeconomic backgrounds, particularly in resource-constrained educational environments. Additionally, while our study provides initial evidence of LingoLift's effectiveness, drawing definitive conclusions about the system's long-term impact requires longitudinal research designs that track children's language development trajectories over extended periods.

Due to equipment constraints described in our study setup, we adopted a one-teacher-one-student design in our field deployment. However, the system architecture supports managing multiple students' personalized learning profiles simultaneously. Future research should investigate how teachers navigate between multiple student profiles in authentic teaching contexts, including potential challenges such as cognitive load during context-switching, risk of confusion, and strategies for adapting and reusing instructional materials across different students.

**Confounding Interventions.** During the study period, children continued to receive other therapeutic interventions such as social skills training, sensory integration therapy, and art education. Although these treatments are not directly related to language development, considering the comprehensive development of cognitive abilities, they may influence LingoLift's specific contribution to the observed learning outcomes.

**Linguistic and Cultural Specificity.** Our language learning architecture incorporates concepts unique to Chinese, such as initials and finals in pinyin instruction, and uses Chinese Hunyuan models to generate culturally relevant content. For different linguistic structural features, lesson generation modules would need to be redesigned and effectiveness validated across different linguistic contexts.

## 8 Conclusion

This paper presents LingoLift, a generative AI-powered system designed to support educators in personalized oral language teaching for autistic children. The system generates teaching materials based on children's interests and dynamically tracks their learning progress, seamlessly integrating these materials into situated learning contexts through AR projection. A 3-week field deployment study with 10 educator-autistic children dyads demonstrated the system's usability, with children showing significant engagement improvements over the study period. LingoLift effectively enhanced traditional teaching practices by streamlining manual processes, enhancing content inspiration, and providing real-time contextual support, while fostering teachers' creative pedagogical adaptations.

Our findings contribute practical insights into the effective application of generative AI in special education, revealing unique requirements of oral language learning for autistic children and critical design considerations for inclusive education contexts.

## Disclosure about Use of LLM

We used GitHub Copilot to assist with code implementation in TypeScript code for project architecture, server setup, and client development. Generative AI did not contribute to the architectural design of the project or to the programming of the children's information database. In the data analysis process, we used *GPT-4o* to assist with Python code for data analysis and visualization in Jupyter Notebooks. However, generative AI was not involved in the formulation of the methods and processes for the data analysis itself. For the prompt engineering in the system, see Appendix C.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) Research Fund for International Scientists (RFIS) (Grant No. W2533160), the Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things (Grant No. 2023B1212010007), the 111 Center (Grant No. D25008), the Guangzhou Municipal Bureau of Education and the Guangzhou Education Foundation (Grant No. 20251213789), and the Department of Education of Guangdong Province (Grant No. SZFNSTJY2025002).

We also thank the **China Computer Federation (CCF) Public Welfare Committee** and the **Tencent Foundation** for recognizing the project and for their encouragement and support.

Special thanks are extended to the teachers and neurodiverse students at **Nansha Qihui School** for their partnership in the deployment study and for their valuable insights that were instrumental to this research.

## References

- [1] Doris A Allen, Isabelle Rapin, and Max Wiznitzer. 1988. Communication disorders of preschool children: the physician's responsibility. *Journal of Developmental & Behavioral Pediatrics* 9, 3 (1988), 164–170.
- [2] Kyle Banker, Douglas Garrett, Peter Bakkum, and Shaun Verch. 2016. *MongoDB in action: covers MongoDB version 3.0*. Simon and Schuster.
- [3] Rongqi Bei, Yajie Liu, Yihe Wang, Yuxuan Huang, Ming Li, Yuhang Zhao, and Xin Tong. 2024. StarRescue: the Design and Evaluation of A Turn-Taking Collaborative Game for Facilitating Autistic Children's Social Skills. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 67, 19 pages. doi:10.1145/3613904.3642829
- [4] L. Belenguer. 2022. AI bias: exploring discriminatory algorithmic decision-making models and the application of possible machine-centric solutions adapted from the pharmaceutical industry. *Ai and Ethics* 2 (2022), 771 – 787. doi:10.1007/s43681-022-00138-8
- [5] Alex Biryukov and Dmitry Khovratovich. 2009. Related-key cryptanalysis of the full AES-192 and AES-256. In *International conference on the theory and application of cryptology and information security*. Springer, 1–18.
- [6] Emily C Bouck. 2009. No Child Left Behind, the Individuals with Disabilities Education Act and functional curricula: A conflict of interest? *Education and Training in Developmental Disabilities* (2009), 3–13.
- [7] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
- [8] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health* 11, 4 (2019), 589–597.
- [9] Yancheng Cao, Yangyang He, Yonglin Chen, Menghan Chen, Shanhe You, Yulin Qiu, Min Liu, Chuan Luo, Chen Zheng, Xin Tong, Jing Liang, and Jiangtao Gong. 2025. Designing LLM-simulated Immersive Spaces to Enhance Autistic Children's Social Affordances Understanding in Traffic Settings. In *Proceedings of the 30th International Conference on Intelligent User Interfaces (IUI '25)*. Association for Computing Machinery, New York, NY, USA, 519–537. doi:10.1145/3708359.3712142
- [10] Sophie Carruthers, A. Pickles, V. Slonims, P. Howlin, and T. Charman. 2020. Beyond intervention into daily life: A systematic review of generalisation following social communication interventions for young children with autism. *Autism Research* 13 (2020), 506 – 522. doi:10.1002/aur.2264
- [11] Sandra Chan, Yiyu Cai, Andrew Lu, Nay Zin Tun, Lihui Huang, and Indhumathi Chandrasekaran. 2016. Virtual reality enhanced pink dolphin game for children with ASD. In *Proceedings of the 3rd Asia-Europe Symposium on Simulation & Serious Gaming*. 215–218.
- [12] Lijia Chen, Pingping Chen, and Zhijian Lin. 2020. Artificial intelligence in education: A review. *IEEE access* 8 (2020), 75264–75278.
- [13] Pawel Chodkiewicz, Pragma Verma, and Grischa Liebel. 2025. *CoVoL: A Cooperative Vocabulary Learning Game for Children with Autism*. Association for Computing Machinery, New York, NY, USA, 1000–1004. https://doi.org/10.1145/3713043.3731521
- [14] Dasom Choi, SoHyun Park, Kyungah Lee, Hwajung Hong, and Young-Ho Kim. 2025. AACessTalk: Fostering Communication between Minimally Verbal Autistic Children and Parents with Contextual Guidance and Card Recommendation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 556, 25 pages. doi:10.1145/3706598.3713792
- [15] Fred D Davis. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly* (1989), 319–340.
- [16] Yao Du and Felix Juefei-Xu. 2023. Generative AI for therapy? Opportunities and barriers for ChatGPT in speech-language therapy. (2023).
- [17] Inge-Marie Eigsti, Ashley B de Marchena, Jillian M Schuh, and Elizabeth Kelley. 2011. Language acquisition in autism spectrum disorders: A developmental review. *Research in autism spectrum disorders* 5, 2 (2011), 681–691.
- [18] Ibrahim El Shemy, Ana Lucia Urrea Echeverria, Gema Erena-Guardia, David Saldaña, Mila Vulchanova, and Michail Giannakos. 2023. Enhancing the vocabulary learning skills of autistic children using augmented reality: a participatory design perspective. In *Proceedings of the 2023 Symposium on Learning, Design and Technology* (Evanston, IL, USA) (LDT '23). Association for Computing Machinery, New York, NY, USA, 87–96. doi:10.1145/3594781.3594791
- [19] Saad Elbeleidy, Terran Mott, Dan Liu, Ellen Do, Elizabeth Reddy, and Tom Williams. 2023. Beyond the session: Centering teleoperators in socially assistive robot-child interactions reveals the bigger picture. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–33.
- [20] Haoxiang Fan, Guanzheng Chen, Xingbo Wang, and Zhenhui Peng. 2024. LessonPlanner: Assisting Novice Teachers to Prepare Pedagogy-Driven Lesson Plans with Large Language Models. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) (UIST '24). Association for Computing Machinery, New York, NY, USA, Article 146, 20 pages. doi:10.1145/3654777.3676390
- [21] Yizhou Fan. [n. d.]. Future Vision and Key Topics of Learning with GenAI: Conceptual Constructions Rooted in Empirical Studies. In *Learning with Generative Artificial Intelligence*. Routledge, 254–260.
- [22] Yizhou Fan. 2025. *Learning with Generative Artificial Intelligence: What Empirical Studies Tell Us*. Taylor & Francis.
- [23] Sue Fletcher-Watson, Alexandra Petrou, Juliet Scott-Barrett, Pamela Dicks, Catherine Graham, Anne O'Hare, Helen Pain, and Helen McConachie. 2016. A trial of an iPad™ intervention targeting social communication skills in children with autism. *Autism* 20, 7 (2016), 771–782.
- [24] Mauricio Fontana De Vargas, Christina Yu, Howard C. Shane, and Karyn Moffatt. 2024. Co-Designing QuickPic: Automated Topic-Specific Communication Boards from Photographs for AAC-Based Language Instruction. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 910, 16 pages. doi:10.1145/3613904.3642080
- [25] Javier Gomez, Letizia Jaccheri, Juan Carlos Torrado, and Germán Montoro. 2018. Leo con lula, introducing global reading methods to children with ASD. In *Proceedings of the 17th ACM Conference on Interaction Design and Children* (Trondheim, Norway) (IDC '18). Association for Computing Machinery, New York, NY, USA, 420–426. doi:10.1145/3202185.3202765
- [26] Tua Haldin-Herrgard. 2000. Difficulties in diffusion of tacit knowledge in organizations. *Journal of Intellectual capital* 1, 4 (2000), 357–365.
- [27] Sangyeal Han and Heetae Yang. 2018. Understanding adoption of intelligent personal assistants: A parasocial relationship perspective. *Industrial Management & Data Systems* 118, 3 (2018), 618–636.
- [28] Holly Hodges, Casey Fealko, and Neelkamal Soares. 2020. Autism spectrum disorder: definition, epidemiology, causes, and clinical evaluation. *Translational pediatrics* 9, Suppl 1 (2020), S55.
- [29] Wayne Holmes, Maya Bialik, and Charles Fadel. 2019. *Artificial intelligence in education promises and implications for teaching and learning*. Center for Curriculum Redesign.

- [30] Patricia Howlin. 1982. Echolalic and spontaneous phrase speech in autistic children. *Journal of Child Psychology and Psychiatry* 23, 3 (1982), 281–293.
- [31] Yiling Hu, Zihong Zhao, and Fang Wen. 2022. Research on the Co-evolution Model of AI Education Ecosystem: The Perspective of Complex System. *Journal of East China Normal University (Educational Sciences)* 40, 9 (2022), 118.
- [32] Kara Hume, Jessica R Steinbrenner, Samuel L Odom, Kristi L Morin, Sallie W Nowell, Brianne Tomaszewski, Susan Szendrey, Nancy S McIntyre, Serife Yücesoy-Özkan, and Melissa N Savage. 2021. Evidence-based practices for children, youth, and young adults with autism: Third generation review. *Journal of autism and developmental disorders* 51, 11 (2021), 4013–4032.
- [33] Neil Humphrey and Wendy Symes. 2013. Inclusive education for pupils with autistic spectrum disorders in secondary mainstream schools: teacher attitudes, experience and knowledge. *International Journal of inclusive education* 17, 1 (2013), 32–46.
- [34] Kasun Indrasiri and Danesh Kuruppu. 2020. *gRPC: up and running: building cloud native applications with Go and Java for Docker and Kubernetes*. O'Reilly Media.
- [35] Rose Iovannone, Glen Dunlap, Heather Huber, and Don Kincaid. 2003. Effective educational practices for students with autism spectrum disorders. *Focus on autism and other developmental disabilities* 18, 3 (2003), 150–165.
- [36] Sinuo Jing, Bozhen Zhu, and Zaiqiao Ye. 2025. Lend A Hand: Designing A Robot for Teaching Social Skills to Children with High-Functioning Autism. In *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction (Melbourne, Australia) (HRI '25)*. IEEE Press, 1368–1372.
- [37] Connie Kasari, Erin Rotheram-Fuller, Jill Locke, and Amanda Gulsrud. 2012. Making the connection: Randomized controlled trial of social skills at school for children with autism spectrum disorders. *Journal of child psychology and psychiatry* 53, 4 (2012), 431–439.
- [38] Connie Kasari, Alexandra Sturm, and Wendy Shih. 2018. SMARTer approach to personalizing intervention for children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research* 61, 11 (2018), 2629–2640.
- [39] Nazilla Khanlou, Attia Khan, Luz Maria Vazquez, and Masood Zangeneh. 2021. Digital literacy, access to technology and inclusion for young adults with developmental disabilities. *Journal of Developmental and Physical Disabilities* 33, 1 (2021), 1–25.
- [40] Hassan Khosravi, Simon Buckingham Shum, Guanliang Chen, Cristina Conati, Yi-Shan Tsai, Judy Kay, Simon Knight, Roberto Martinez-Maldonado, Shazia Sadiq, and Dragan Gašević. 2022. Explainable artificial intelligence in education. *Computers and education: artificial intelligence* 3 (2022), 100074.
- [41] Kamran Khowaja, Dena Al-Thani, Asma Osman Hassan, Asadullah Shah, and Siti Salwah Salim. 2020. Mobile augmented reality app for children with autism spectrum disorder (ASD) to learn vocabulary (MARVoc): from the requirement gathering to its initial evaluation. In *International Conference on Human-Computer Interaction*. Springer, 424–437.
- [42] Kamran Khowaja, Bilikis Banire, Dena Al-Thani, Mohammed Tahri Sqalli, Aboubakr Aqle, Asadullah Shah, and S. Salim. 2020. Augmented Reality for Learning of Children and Adolescents With Autism Spectrum Disorder (ASD): A Systematic Review. *IEEE Access* 8 (2020), 78779–78807. doi:10.1109/access.2020.2986608
- [43] Kamran Khowaja and Siti Salwah Salim. 2019. Serious game for children with autism to learn vocabulary: an experimental evaluation. *International journal of human-computer interaction* 35, 1 (2019), 1–26.
- [44] Kamran Khowaja and Siti Salwah Salim. 2020. A framework to design vocabulary-based serious games for children with autism spectrum disorder (ASD). *Universal Access in the Information Society* 19, 4 (2020), 739–781.
- [45] Robert L Koegel, Mary C O'dell, and Lynn Kern Koegel. 1987. A natural language teaching paradigm for nonverbal autistic children. *Journal of autism and developmental disorders* 17, 2 (1987), 187–200.
- [46] Sofia Kotsi, Spyridoula Handrinou, Georgia Iatraki, and Spyridon-Georgios Soulis. 2025. A review of artificial intelligence interventions for students with autism spectrum disorder. *Disabilities* 5, 1 (2025), 7.
- [47] Stephen Krashen. 1981. Second language acquisition. *Second Language Learning* 3, 7 (1981), 19–39.
- [48] David R Krathwohl. 2002. A revision of Bloom's taxonomy: An overview. *Theory into practice* 41, 4 (2002), 212–218.
- [49] Yu-Ju Lan. 2020. Immersion into virtual reality for language learning. In *Psychology of learning and motivation*. Vol. 72. Elsevier, 1–26.
- [50] Jean Lave and Etienne Wenger. 1991. *Situated learning: Legitimate peripheral participation*. Cambridge university press.
- [51] Jungeun Lee, Suwon Yoon, Kyoosik Lee, Eunae Jeong, Jae-Eun Cho, Wonjeong Park, Dongsun Yim, and Inseok Hwang. 2024. Open Sesame? Open Salami! Personalizing Vocabulary Assessment-Intervention for Children via Pervasive Profiling and Bespoke Storybook Generation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 120, 32 pages. doi:10.1145/3613904.3642580
- [52] Aaleyah Lewis, Aayushi Dangol, Hyewon Suh, Abbie Olszewski, James Fogarty, and Julie A. Kientz. 2025. Exploring AI-Based Support in Speech-Language Pathology for Culturally and Linguistically Diverse Children. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 558, 19 pages. doi:10.1145/3706598.3714131
- [53] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems* 33 (2020), 9459–9474.
- [54] Guang Li, Mohammad Amin Zarei, Goudarz Alibakhshi, and Akram Labbafi. 2024. Teachers and educators' experiences and perceptions of artificial-powered interventions for autism groups. *BMC psychology* 12, 1 (2024), 199.
- [55] Yujia Liu, Siyu Zha, Yuewen Zhang, Yanjin Wang, Yangming Zhang, Qi Xin, Lun Yiu Nie, Chao Zhang, and Yingqing Xu. 2025. BrickSmart: Leveraging Generative AI to Support Children's Spatial Language Learning in Family Block Play. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 645, 19 pages. doi:10.1145/3706598.3714212
- [56] Catherine Lord. 1997. Language and communication in autism. *Handbook of autism and pervasive developmental disorders (1997)*, 195–225.
- [57] O Ivar Lovaas and Tristram Smith. 2003. Early and intensive behavioral intervention in autism. *Evidence-based psychotherapies for children and adolescents* (2003), 325–340.
- [58] Camillo Lugaresi, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo-Ling Chang, Ming Guang Yong, Juhyun Lee, et al. 2019. Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172* (2019).
- [59] Sue Lyle. 2008. Dialogic teaching: Discussing theoretical contexts and reviewing evidence from classroom practice. *Language and education* 22, 3 (2008), 222–240.
- [60] Yue Lyu, Di Liu, Pengcheng An, Xin Tong, Huan Zhang, Keiko Katsuragawa, and Jian Zhao. 2024. EMooly: Supporting Autistic Children in Collaborative Social-Emotional Learning with Caregiver Participation through Interactive AI-infused and AR Activities. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 4, Article 203 (Nov. 2024), 36 pages. doi:10.1145/3699738
- [61] Johnny L Matson, Robert D Rieske, and Lindsey W Williams. 2013. The relationship between autism spectrum disorders and attention-deficit/hyperactivity disorder: An overview. *Research in developmental disabilities* 34, 9 (2013), 2475–2484.
- [62] Vânia Mendonça, Luísa Coheur, and Alberto Sardinha. 2015. VITHEA-Kids: a Platform for Improving Language Skills of Children with Autism Spectrum Disorder. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility (Lisbon, Portugal) (ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 345–346. doi:10.1145/2700648.2811371
- [63] Reza Hadi Mogavi, Chao Deng, Justin Juho Kim, Pengyuan Zhou, Young D Kwon, Ahmed Hosny Saleh Metwally, Ahmed Tlili, Simone Bassanelli, Antonio Bucchiarone, Sujit Gujar, et al. 2023. Exploring user perspectives on chatgpt: Applications, perceptions, and implications for ai-integrated education. *arXiv preprint arXiv:2305.13114* (2023).
- [64] Maggie A. Mosher, Adam C. Carreon, S. Craig, and Lindsay Ruhter. 2021. Immersive Technology to Teach Social Skills to Students with Autism Spectrum Disorder: a Literature Review. *Review Journal of Autism and Developmental Disorders* 9 (2021), 334–350. doi:10.1007/s40489-021-00259-6
- [65] L Mottron and J Burack. 2006. Autism: A different perception-Preface. 3 pages.
- [66] Ian SP Nation and ISP Nation. 2001. *Learning vocabulary in another language*. Vol. 10. Cambridge university press Cambridge.
- [67] Leslie C. Neely, J. Ganz, John L. Davis, Margot B. Boles, E. Hong, Jennifer Ninci, and Whitney D. Gilliland. 2016. Generalization and Maintenance of Functional Living Skills for Individuals with Autism Spectrum Disorder: a Review and Meta-Analysis. *Review Journal of Autism and Developmental Disorders* 3 (2016), 37–47. doi:10.1007/s40489-015-0064-7
- [68] David Nunan. 2005. An introduction to task-based language teaching (Nunan, 2004). *The Asian EFL Journal Quarterly June 2005 Volume 7* (2005), 25–28.
- [69] Samuel L Odom, William H Brown, Timothy Frey, Necdet Karasu, Lora Lee Smith-Canter, and Phillip S Strain. 2003. Evidence-based practices for young children with autism: Contributions for single-subject design research. *Focus on autism and other developmental disabilities* 18, 3 (2003), 166–175.
- [70] Julia Parish-Morris, Elizabeth A Hennon, Kathy Hirsh-Pasek, Roberta Michnick Golinkoff, and Helen Tager-Flusberg. 2007. Children with autism illuminate the role of social intention in word learning. *Child development* 78, 4 (2007), 1265–1287.
- [71] Jungin Park, Gahyeon Bae, Jueon Park, Seo Kyoung Park, Yeon Soo Kim, and Sangsu Lee. 2023. AEDLE: Designing Drama Therapy Interface for Improving Pragmatic Language Skills of Children with Autism Spectrum Disorder Using AR. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (Hamburg, Germany) (CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 25, 7 pages. doi:10.1145/3544549.3585809

- [72] Lauren Parsons, Reinie Cordier, Natalie Munro, Annette Joosten, and Renee Speyer. 2017. A systematic review of pragmatic language interventions for children with autism spectrum disorder. *PLoS one* 12, 4 (2017), e0172242.
- [73] James W Partington. 2008. *The assessment of basic language and learning skills-revised (the ABLLS-R)*. Behavior Analysts.
- [74] Kaska Porayska-Pomsta, Alyssa M Alcorn, Katerina Avramides, Sandra Beale, Sara Bernardini, Mary Ellen Foster, Christopher Frauenberger, Judith Good, Karen Gulberg, Wendy Keay-Bright, et al. 2018. Blending human and artificial intelligence to support autistic children's social communication skills. *ACM Transactions on Computer-Human Interaction (TOCHI)* 25, 6 (2018), 1–35.
- [75] Johanna Pöysä-Tarhonen, Nafisa Awwal, Päivi Häkkinen, and Suzanne Otieno. 2021. Joint attention behaviour in remote collaborative problem solving: exploring different attentional levels in dyadic interaction. *Research and Practice in Technology Enhanced Learning* 16, 1 (2021), 11.
- [76] Barry M Prizant, Amy M Wetherby, Emily Rubin, Amy C Laurent, and Patrick J Rydell. 2006. *The SCERTS model: A comprehensive educational approach for children with autism spectrum disorders, Vol. 1*. Paul H. Brookes Publishing Co.
- [77] Md Mostafizer Rahman and Yutaka Watanobe. 2023. ChatGPT for education and research: Opportunities, threats, and strategies. *Applied sciences* 13, 9 (2023), 5783.
- [78] Rishabh Ranawat, Ashwin Venkataraman, and Lakshminarayanan Subramanian. 2021. CollectiveTeach: A System To Generate And Sequence Web-Annotated Lesson Plans. In *Proceedings of the 4th ACM SIGCAS Conference on Computing and Sustainable Societies (Virtual Event, Australia) (COMPASS '21)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3460112.3471938
- [79] Isabelle Rapin and Michelle Dunn. 1997. Language disorders in children with autism. In *Seminars in pediatric neurology, Vol. 4*. Elsevier, 86–92.
- [80] Jack C Richards. 2001. *Curriculum development in language teaching*. Cambridge university press.
- [81] Vladimir E Robles-Bykbaev, Martín López-Nores, José J Pazos-Arias, and Daysi Arévalo-Lucero. 2015. SPELTA: An expert system to generate therapy plans for speech and language disorders. *Expert Systems with Applications* 42, 21 (2015), 7641–7651.
- [82] Lisa Ruble and John H McGrew. 2013. Teacher and child predictors of achieving IEP goals of children with autism. *Journal of autism and developmental disorders* 43, 12 (2013), 2748–2763.
- [83] Lisa A Ruble and Nancy J Dalrymple. 2002. Compass: A parent-Teacher collaborative model for students with autism. *Focus on autism and other developmental disabilities* 17, 2 (2002), 76–83.
- [84] Lisa A Ruble, John H McGrew, Michael Toland, Nancy Dalrymple, Medina Adams, and Claire Snell-Rood. 2018. Randomized control trial of COMPASS for improving transition outcomes of students with autism spectrum disorder. *Journal of Autism and Developmental Disorders* 48, 10 (2018), 3586–3595.
- [85] Brenda Scheuermann and Jo Webber. 2002. Autism: Teaching does make a difference. (No Title) (2002).
- [86] Bertrand Schneider. 2015. *The Role of Joint Visual Attention in Collaborative Learning*. Stanford University.
- [87] Donghee Shin. 2021. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International journal of human-computer studies* 146 (2021), 102551.
- [88] Donghoon Shin, Jaeyoon Song, Seokwoo Song, Jisoo Park, Joonhwan Lee, and Soojin Jun. 2020. TalkingBoogie: collaborative mobile AAC system for non-verbal children with developmental disabilities and their caregivers. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [89] Richard L Simpson. 2005. Evidence-based practices and students with autism spectrum disorders. *Focus on autism and other developmental disabilities* 20, 3 (2005), 140–149.
- [90] Daniel L Stufflebeam. 2000. The CIPP model for evaluation. In *Evaluation models: Viewpoints on educational and human services evaluation*. Springer, 279–317.
- [91] Hyewon Suh, Aayushi Dangol, Hedda Meadan, Carol A. Miller, and Julie A. Kientz. 2024. Opportunities and Challenges for AI-Based Support for Speech-Language Pathologists. In *Proceedings of the 3rd Annual Meeting of the Symposium on Human-Computer Interaction for Work* (Newcastle upon Tyne, United Kingdom) (CHIWORK '24). Association for Computing Machinery, New York, NY, USA, Article 14, 14 pages. doi:10.1145/3663384.3663387
- [92] Xingwu Sun, Yanfeng Chen, Yiqing Huang, Ruobing Xie, Jiaqi Zhu, Kai Zhang, Shuaipeng Li, Zhen Yang, Jonny Han, Xiaobo Shu, Jiahao Bu, Zhongzhi Chen, Xuemeng Huang, Feng Lian, Saiyong Yang, Jianfeng Yan, Yuyuan Zeng, Xiaoqing Ren, Chao Yu, Lulu Wu, Yue Mao, Jun Xia, Tao Yang, Suncong Zheng, Kan Wu, Dian Jiao, Jinbao Xue, Xipeng Zhang, Decheng Wu, Kai Liu, Dengpeng Wu, Guanghui Xu, Shaohua Chen, Shuang Chen, Xiaowei Feng, Yigeng Hong, Junqiang Zheng, Chengcheng Xu, Zong-Rui Li, Xi Kuang, Jian hua Hu, Yiqi Chen, Yuchi Deng, Guiyang Li, Ao Liu, Chenchen Zhang, Shi-He Hu, Zilong Zhao, Zi-Hao Wu, Yao Ding, Weichao Wang, Han Liu, Roberts Wang, Haoyang Fei, Peijie Yu, Ze Zhao, Xun Cao, Hai Wang, Fusheng Xiang, Meng-Sheng Huang, Zhiyu Xiong, Bin Hu, Xue yan Hou, Lei Jiang, Jia bing Ma, Jiajia Wu, Yaping Deng, Yi Shen, Qian Wang, Weijie Liu, Jie Liu, Meng Chen, Liang Dong, Wei Jia, Hu Chen, Feifei Liu, Ruixin Yuan, Huilin Xu, Zhenxiang Yan, Tengfei Cao, Zhichao Hu, Xinhua Feng, Dong Du, Ting-Ting Yu, Yang-Dan Tao, Feng Zhang, Jianchen Zhu, Chengzhong Xu, Xirui Li, Chong Zha, Ouyang Wen, Yi Xia, Xiang Li, Ze He, Rongpeng Chen, Jiawei Song, Ruibin Chen, Fan Jiang, Chongqing Zhao, and Bo Wang. 2024. Hunyuan-Large: An Open-Source MoE Model with 52 Billion Activated Parameters by Tencent. *ArXiv abs/2411.02265* (2024). doi:10.48550/arXiv.2411.02265
- [93] Mark L Sundberg. 2008. *VB-MAPP Verbal Behavior Milestones Assessment and Placement Program: a language and social skills assessment program for children with autism or other developmental disabilities: guide*. Mark Sundberg.
- [94] Mark L Sundberg. 2008. *VB-MAPP Verbal Behavior Milestones Assessment and Placement Program: a language and social skills assessment program for children with autism or other developmental disabilities: guide*. Mark Sundberg.
- [95] Helen Tager-Flusberg and Connie Kasari. 2013. Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism research* 6, 6 (2013), 468–478.
- [96] Tiffany Y. Tang, Jiasheng Xu, and Pinata Winoto. 2019. An Augmented Reality-Based Word-Learning Mobile Application for Children with Autism to Support Learning Anywhere and Anytime: Object Recognition Based on Deep Learning. In *Universal Access in Human-Computer Interaction. Multimodality and Assistive Environments: 13th International Conference, UAHCI 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part II* (Orlando, FL, USA). Springer-Verlag, Berlin, Heidelberg, 182–192. doi:10.1007/978-3-030-23563-5\_16
- [97] Yilin Tang, Liuqing Chen, Ziyu Chen, Wenkai Chen, Yu Cai, Yao Du, Fan Yang, and Lingyun Sun. 2024. EmoEden: Applying Generative Artificial Intelligence to Emotional Learning for Children with High-Function Autism. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 1001, 20 pages. doi:10.1145/3613904.3642899
- [98] Tencent Hunyuan Team, Ao Liu, Botong Zhou, Can Xu, Chayse Zhou, ChenChen Zhang, Chengcheng Xu, Chenhao Wang, Decheng Wu, Dengpeng Wu, et al. 2025. Hunyuan-TurboS: Advancing Large Language Models through Mamba-Transformer Synergy and Adaptive Chain-of-Thought. *arXiv preprint arXiv:2505.15431* (2025).
- [99] Rianne Verschuur, Robert Didden, Russell Lang, Jeff Sigafos, and Bibi Huskens. 2014. Pivotal response treatment for children with autism spectrum disorders: A systematic review. *Review Journal of Autism and Developmental Disorders* 1, 1 (2014), 34–61.
- [100] Laurie A Vismara and Gregory L Lyons. 2007. Using perseverative interests to elicit joint attention behaviors in young children with autism: Theoretical and clinical implications for understanding motivation. *Journal of Positive Behavior Interventions* 9, 4 (2007), 214–228.
- [101] Pinata Winoto and Tiffany Y. Tang. 2018. Two Lightweight and Customizable Picture-based Word-Learning Mobile Applications for Chinese Children with Autism. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (Singapore, Singapore) (UbiComp '18). Association for Computing Machinery, New York, NY, USA, 291–294. doi:10.1145/3267305.3267628
- [102] Mary Ann Winter-Messiers. 2007. From tarantulas to toilet brushes: Understanding the special interest areas of children and youth with Asperger syndrome. *Remedial and Special Education* 28, 3 (2007), 140–152.
- [103] Jiazhou Wu, Min Fan, and Liyan Sheng. 2025. Co-designing and evaluating virtual tutors in a social language learning app for children with autism spectrum disorder. *CoDesign* (2025), 1–28.
- [104] Xingrui Xie, Han Liu, Wenzhe Hou, and Hongbin Huang. 2023. A brief survey of vector databases. In *2023 9th International Conference on Big Data and Information Analytics (BigDIA)*. IEEE, 364–371.
- [105] Weijian Yan, Victoria Lynn Lowell, and Li Yang. 2024. Developing English language learners' speaking skills through applying a situated learning approach in VR-enhanced learning experiences. *Virtual Reality* 28, 4 (2024), 167.
- [106] Oguzcan Yildirim and Elif Surer. 2021. Developing adaptive serious games for children with specific learning difficulties: A two-phase usability and technology acceptance study. *JMIR Serious Games* 9, 2 (2021), e25997.
- [107] Chao Zhang, Cheng Yao, Jiayi Wu, Weijia Lin, Lijuan Liu, Ge Yan, and Fangtian Ying. 2022. StoryDrawer: A Child-AI Collaborative Drawing System to Support Children's Creative Visual Storytelling. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 311, 15 pages. doi:10.1145/3491102.3501914
- [108] Zheng Zhang, Ying Xu, Yanhao Wang, Bingsheng Yao, Daniel Ritchie, Tongshuang Wu, Mo Yu, Dakuo Wang, and Toby Jia-Jun Li. 2022. StoryBuddy: A Human-AI Collaborative Chatbot for Parent-Child Interactive Storytelling with Flexible Parental Involvement. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 218, 21 pages. doi:10.1145/3491102.3517479

- [109] Leming Zhou, Jie Bao, I Made Agus Setiawan, Andi Saptono, and Bambang Parmanto. 2019. The mHealth app usability questionnaire (MAUQ): development and validation study. *JMIR mHealth and uHealth* 7, 4 (2019), e11500.
- [110] V. Žalys. 2021. Interactive Multi-Sensory Environments for Children with Autism Spectrum Disorders. *ATEE 2020 - Winter Conference. Teacher Education for Promoting Well-Being in School. Suceava, 2020* (2021). doi:10.18662/lumproc/atee2020/39

## Appendices

### A Post-lesson Questionnaire

ID	Item
<b>(1) Personalized Alignment</b>	
<i>Scale: 1=Strongly Disagree, 4=I'm not sure, 7=Strongly Agree</i>	
Q1	LingoLift recommended appropriate personalized learning objectives for the child.
Q3	LingoLift recommended appropriate personalized learning themes for the child.
<b>(2) Lesson Coherence</b>	
<i>Scale: 1=Strongly Disagree, 7=Strongly Agree</i>	
Q2	The generated teaching content is well-aligned with the learning objectives.
Q4	<b>[Within-lesson]</b> The teaching content across different modules (articulation, tact, linguistic, intraverbal) maintains good coherence and thematic consistency.
Q5	<b>[Across-lesson]</b> The teaching content of this lesson maintains good coherence with previous lessons.
<b>(3) AI-Generated Content Satisfaction</b>	
<i>Scale: 1=Strongly Disagree, 4=I'm not sure, 7=Strongly Agree</i>	
Q6	I am satisfied with the content and quality of the generated text teaching materials.
Q7	I am satisfied with the content and quality of the generated image teaching materials.
Q8	Overall, I am satisfied with the teaching content generated by LingoLift for this lesson.
Q10	During the lesson, LingoLift's lesson plan provided me with effective teaching guidance.
<b>(4) Engagement</b>	
<i>Scale: 1=Strongly Disagree, 4=I'm not sure, 7=Strongly Agree</i>	
Q11	The child's engagement and focus improved during the lesson.
<b>(5) Learning Outcomes</b>	
<i>Scale: 1=No improvement, 4=Moderate improvement, 7=Significant improvement</i>	
Q12	Degree of improvement in the child's pronunciation during this lesson.
Q13	Degree of improvement in the child's naming skills during this lesson.
Q14	Degree of improvement in the child's language structure skills during this lesson.
Q15	Degree of improvement in the child's conversation skills during this lesson.
<i>Scale: 1=Strongly Disagree, 7=Strongly Agree</i>	
Q16	Overall, the child achieved the preset speech teaching objectives for this lesson.
<b>Preparation Time</b>	
Q9	Approximately how much time did you spend preparing for this lesson using LingoLift?

## B Post-study Questionnaire

---

ID	Item
Q1	This system provides an acceptable way to deliver speech teaching services.
Q2	This system has improved the way I deliver teaching services.
Q3	This system helps me effectively manage the children's status.
Q4	I feel comfortable using this application to interact with children.
Q5	This system helps me get my work done faster.
Q6	This system enhances my work performance.
Q7	This system improves my productivity during work.
Q8	This system makes me more efficient at work.
Q9	This system makes my work easier.
Q10	I believe this system is useful in my work.
Q11	Learning this system is easy for me.
Q12	I can easily make this system work the way I want.
Q13	I think this system is easy to understand.
Q14	It is easy for me to become skilled at using this system.
Q15	I think this system is easy to use.
Q16	I would like to use this app again.
Q17	Overall, I'm satisfied with this system.

---

## C Examples of Prompt Engineering in the Project (TypeScript)

```

/**
 * Example: Prompt used for generating a lesson plan
 */
const promptLesson = `This is the teaching content generation module for this stage:
please provide a complete teaching plan that can be directly wrapped as a string
inside my React Native Text component, using \n for line breaks.
Do not include Text, jsx, or any irrelevant content in the returned result.
The length should be around 300 Chinese characters, with professional wording.
You are an autism education expert in China, preparing a teaching module
for the VB-MAPP naming domain for children with autism.
The teaching objective is: |${safeDescriptions}|.
The teaching scenario is: |${learningGoals?.TeachingScene?.major}| - |${learningGoals?.TeachingScene?.activity}|.
The vocabulary you can use includes: |${cardsContent}|.
Please generate a concrete teaching procedure that connects these words
and the scenario as much as possible, while aligning with the naming domain objective.
Directly output the content, without any extra explanation!`;

/**
 * Example: Prompt used for generating words of Cards
 */
const promptCards = `Please return a string: you must strictly use Mandarin pinyin consonant(s) corresponding to |${selectedPinyin}|.
Our teaching scenario is |${Goals?.TeachingScene?.major}|, |${Goals?.TeachingScene?.activity}|.
The words must be thematically relevant and non-repetitive.

Directly give me a string formatted as an array wrapped in [ ], not in JSON.
The array must contain exactly four words. Each word should be:
1. A meaningful Chinese word (Daily Life, suitable for children),
2. Written in characters followed by its pinyin in parentheses,
3. Separated by English commas.

All four words must begin with |${selectedPinyin}| and be contextually appropriate.`;

console.log(promptCards);
// Example outcome: [high-speed rail(gao tie), juice(guo zhi), park(gong yuan), bear(gou xiong)]

/**
 * Example: Prompt used for generating child-centered learning scenarios
 */
const promptScenario = `You are an experienced children's education expert.
Based on the following reinforcers, please generate the names of three
child-centered teaching scenarios that are engaging and motivational
for children with autism. Each scenario must integrate the reinforcer
in parentheses after the title.

Output format:
['Scenario A (Reinforcer1)', 'Scenario B (Reinforcer2)', 'Scenario C (Reinforcer3)'].

Reinforcer list: |${reinArr.join(', ')}|.`;

/**
 * Example: Text-to-Image prompt for generating child-friendly illustrations
 */
const promptImage = `You are an illustrator specializing in autism-friendly drawings.
Please create an illustration of the object: |${item.value}|,
using the style: |${imageStyle}| (either "realistic" or "cartoon").

Constraints:
1. Must be suitable for children: no NSFW, violent, bloody, scary, or age-inappropriate content.
2. Autism-friendly: soft lines, calm colors, minimal clutter, clear focus.
3. Must fit the scenario: |${learningGoals?.TeachingScene?.major}|.
4. Object |${item.value}| should be clear, positive, and inviting.
5. Background supports but does not distract.
6. Consistent with professional educational materials for autism.`;

```

## D Semi-structured Interview Guide for Formative Study

### D.1 Understanding Children and Teaching Context

- (1) Could you briefly introduce the children in the videos and their language assessment results?
- (2) How do you adapt your teaching approach for different students? What factors do you consider?
- (3) How do you determine and use personalized reinforcers in teaching?
- (4) Do you use interactive devices in your classroom? How do they help?

## D.2 Validating Observations

- (1) We identified these ASD characteristics from the videos. Are our observations accurate?
  - Inappropriate behaviors, non-compliance with instructions
  - Communication deficits, dependence on reinforcers
  - Attention issues, lack of eye contact
  - Pronunciation difficulties, repetitive behaviors
- (2) We summarized these personalized teaching factors. What's your feedback?
  - Reinforcer preferences and language ability levels
  - Instruction compliance and complexity tolerance
  - Learning styles and age considerations
- (3) Would connecting teaching materials to real-life contexts improve effectiveness?
- (4) What are your main pain points in lesson preparation and teaching?

## D.3 AI Tool Design Feedback

- (1) What technological opportunities could improve your teaching process?
- (2) Should we maintain content consistency across different learning modules (pronunciation → naming → dialogue)?
- (3) Would interactive projection technology be useful compared to traditional cards?
- (4) Are our personalized content generation factors reasonable?
- (5) Would desktop projection better suit children's visual habits or worsen eye contact issues?

## E Semi-structured Interview Guide for Field Deployment Study

### E.1 Satisfaction and Overall Experience

- (1) Based on your questionnaire responses, could you explain why you were satisfied/dissatisfied with specific aspects?
  - Are you satisfied with LingoLift's overall system design?
  - Are you satisfied with the quality of generated personalized learning materials and lesson plans?

### E.2 Personalization Features

- (1) Based on your personalization ratings, why do you think LingoLift achieved personalization?
  - In which aspects does LingoLift best demonstrate personalized consideration for children?
  - How does the system's ability to generate lesson plans based on teacher-filled assessment information help create appropriate learning goals?
  - How do teaching themes based on children's reinforcers help engage student interest?
- (2) What aspects of personalized teaching need improvement? What personalized features would you like to see added in the future?
  - Do you think more interactive activities could be designed to improve student engagement?
  - Would you like to see reward mechanisms or personalized mini-games integrated into the teaching process?
  - How could the system better support attention and focus during lessons?

### E.3 Teaching Preparation Support

- (1) How did LingoLift support your lesson preparation process?
- (2) How did your preparation process change compared to traditional lesson planning? How did GenAI change your preparation workflow?
  - How does the system's material generation compare to traditional online material searching?
  - How accurate are the generated pronunciation materials and visual resources?
  - Do you follow AI-generated lesson structures, or do you customize them significantly?
- (3) Do you primarily use custom teaching goals or system-recommended ones?

### E.4 In-class Teaching Support

- (1) How did LingoLift support your actual teaching process?
- (2) How did it change your traditional teaching approach? In what aspects did you see improvements (engagement, learning outcomes)?
  - How effective were interactive features like movable images for teaching spatial concepts?
  - What additional interactive capabilities would you like to see (e.g., AR glasses, scenario-based activities)?

### E.5 System Functionality and Usability

- (1) What difficulties did you encounter while using the system for teaching?
- (2) Which functions did you find most helpful? Why?

- (3) Which functions do you think are redundant or need improvement (interface, interaction, content selection)?

## E.6 Learning Outcomes and Professional Accuracy

- (1) Have you noticed improvements in children’s learning speed compared to before?
- (2) Do you think the system has sufficient accuracy and professionalism? Does the generated content demonstrate good accuracy and expertise?
- (3) Do you think the system is suitable for novice teachers or parents to operate?

## F Thematic Teaching Design Reference

### F.1 Velar Sound-Related Thematic Teaching

**Table 3: Velar sound teaching themes.**

Teaching Theme	Teaching Content	Teaching Objectives
Let’s Challenge Together	Tongue perception and motor training; spinning-wheel games to make basic speech practice engaging; phoneme induction training.	Improve students’ tongue-root lifting ability and enhance initiative in individualized practice.
Interesting Sounds	Gamified situational group teaching; simulate animal sounds and actions; give new names to pronunciations; velar phoneme /g/ acquisition.	Mobilize multi-sensory participation and strengthen imitation abilities through playful tasks.
I’m the Class Helper	Velar /g/ sound–rhyme combination training; vocabulary learning of “switch,” “drink water,” “thirsty”; generalization in daily routines.	Provide more oral expression opportunities via class roles and promote transfer of /g/ and /k/ sounds.
I Can Host Guests	Little Bear’s Guest storybook learning; “today I’m the host” role-play; class snack sharing; anti-nasalization training.	Create interpersonal interaction scenarios and improve accurate production of velar-related words.
Supermarket Shopping	Create a mini supermarket; learn shopping steps; in-class and home/community practice.	Consolidate simple sentence expression with velar sounds and practice situational dialogue to express needs.
Beautiful Campus	Campus tree-planting scenario; serve as campus tour guide; maintain campus together.	Use simple sentences to introduce objects and master tool-usage sentences during labor activities.

## F.2 Alveolar Sound Training Themes

**Table 4: Alveolar sound teaching themes.**

Teaching Theme	Teaching Content	Teaching Purpose
Interesting Sounds	Design guessing games and sound games to make basic speech skill training fun; phoneme /d/ induction training; sound-rhyme combination training.	Promote tongue-tip lifting ability and enable expression of /d/-related syllables/characters.
What to Do When I Grow Up	Recognize professions; “what to do when I grow up.”	Recognize common /d/-related professions in scenarios and establish appropriate classroom career concepts.
Diversity of Professions	Profession guessing; “how much do you know about professions”; initial profession experience.	Consolidate expression ability of words related to consonant /d/ and experience profession diversity in group activities.
I’m a Civilized Student	Learn civilized etiquette; recognize teachers; learn greeting methods; judge right and wrong in daily behaviors.	Guide students to actively greet others and provide guidance for daily campus/class behaviors using positive examples.
New Friends: Nobita and Doraemon	Self-introduction; visit Nobita’s home; treasure-hunting game; sports activities.	Consolidate /d/ pronunciation in self-introduction and home-visit segments; improve pragmatic and dialogue abilities.
I Love Nature	Happy farm; forest kingdom; get close to nature; read “I Love Nature” storybook.	Name animals and plants in nature; cultivate nature awareness and encourage care and environmental protection.