Exploring Parent Involvement in E-book Joint Reading with Voice Agents

Daniel Vargas-Diaz[a,*], Jisun Kim[b], Sulakna Karunaratna[a], Caroline Byrd Hornburg[b], Koeun Choi[b], Sang Won Lee[a]

[a]*Department of Computer Science, Virginia Tech, 24061 VA, Blacksburg, USA*
[b]*Department of Human Development and Family Science, Virginia Tech, 24061 VA, Blacksburg, USA*

Abstract

Parent-child interactions during joint reading play an important role in young children's cognitive and language development. However, contemporary digital book formats—such as e-books or audiobooks—often overlook the role of the parent in reading the text by either dubbing voice narration over it or reading it aloud automatically. With the advancement and prevalence of voice-based conversational artificial intelligence (AI) agents, AI reading an e-book emerges as a novel reading experience, yet reducing the role of parents during read-alouds similarly. When the reading experience becomes less of a joint activity between children and parents, the potential benefits children can gain from reading may diminish. In this user study involving 11 parent-child pairs, we aimed to explore how voice agents (VAs) could be used to create an interactive digital space to 1) promote parental engagement in joint e-book reading with children and 2) enhance parents' and children's joint reading experiences. We developed and evaluated *TaleMate*, an interactive joint reading app that allows parents and children ages 3–6 years to assign different AI voices to the characters from a book while enabling parents to embody one of the characters to read the book with the voice agents. We found that the system supported children's engagement and story recall. Parents reported that they found value in the interactivity of the system and enjoyed a participatory, joint reading experience where both they and their children could choose which characters to embody. These findings offer insights into design considerations for researchers interested in developing applications that facilitate collaborative reading experiences involving parents, children, and voice agents.

Keywords: voice agents, joint reading, e-books, parents, children

1. Introduction

Joint reading refers to a collective activity where an adult, such as a parent or caregiver, and a child focus together on the same words and content (Zivan and Horowitz-Kraus, 2020). In a joint reading setting, a parent and a child either read aloud together or the parent reads

*Corresponding author
Email addresses: danielvargasdiaz@vt.edu (Daniel Vargas-Diaz), jisunk@vt.edu (Jisun Kim), sulaknak@vt.edu (Sulakna Karunaratna), chornburg@vt.edu (Caroline Byrd Hornburg), koeun@vt.edu (Koeun Choi), sangwonlee@vt.edu (Sang Won Lee)

aloud to the child (Kunkun Zhang and Torr, 2016). Read-alouds are common and effective strategies employed during joint reading (Lennox, 2013; Schmidt et al., 2023). Such parent-child reading practices are crucial in fostering young children's vocabulary, comprehension, and critical thinking skills, as demonstrated by empirical evidence (Bus et al., 1995; Miranda-Cueva and Cabanillas-Carbonell, 2020; Krcmar and Cingel, 2014; Wesseling et al., 2017).

However, the advancement of digital technologies has significantly transformed traditional reading practices, introducing new formats such as audiobooks and electronic books (e-books) with voiceover (Larson, 2015). Despite the well-documented benefits of joint reading (Ewin et al., 2021; Kucirkova and Grøver, 2024), the emergence of e-book formats featuring dubbed or automated reading may limit opportunities for parental involvement by reducing their engagement with the text and content of the books.

This challenge has been documented in previous research on mothers of young children, which reveals that electronic devices are rarely used for joint reading compared to traditional printed texts (Nicholas and Paatsch, 2021). A large-scale survey of U.S. parents of children aged 3 to 6 years found that while many parents (49%) allow their children to use e-books, only a small subset (12% of the total) engages with their children while using e-books (Etta, 2019). This trend may increasingly prevent young children from experiencing the benefits of joint reading with parents.

A recent survey compared the effects of e-books and printed books, highlighting the importance of content-relevant design features in e-books (Furenes et al., 2021). The analysis found that when e-books incorporated content-related enhancements that allowed children to interact and gain information about the story, they outperformed print books, particularly when parents were involved. These findings suggest that the design of e-books with voiceovers may have overlooked the opportunity for parents to be actively engaged in the reading process with their children. However, the conventional design of narrated e-books, which uses a pre-recorded linear voiceover, may leave limited room for parents to interact with the text or content during the narration.

Recent explorations in human-computer interaction have proposed using voice agents (VAs) as reading partners, rather than sole readers, thus introducing novel, interactive, and conversational dimensions to the e-book reading experience and addressing some of the format's inherent limitations (Marchetti and Valente, 2018). In some cases, VAs take on most of the responsibility for reading, including reading aloud and asking questions (Xu and Warschauer, 2020). On the other hand, there have been efforts to include parents in the reading process. In most cases, however, the role of VAs is rather that of a facilitator in parent-child joint reading, primarily asking questions about the content based on what they listen to (Lin et al., 2022; Zhang et al., 2022; Xu et al., 2023).

In our work, we focus on e-books with voice-over, where a voice agent acts as a reader who reads aloud a storybook. Despite the educational benefits of read-aloud during joint reading for young children (Lennox, 2013; Schmidt et al., 2023), the potential for integrating VAs into shared, interactive read-aloud involving both parents and children remains largely underexplored. Therefore, the current study aimed to address this gap by exploring the concept of *parent-child-VA reading*, which engages parents in joint reading experiences symbiotically with VAs. Unlike traditional e-books and audiobooks that offer linear voice narration, TaleMate allows both parents and children to actively participate by choosing roles within the story. This feature enables users not only to assign specific characters to VAs but also to take on characters themselves, allowing them to engage more fully in the narrative as active participants. The involvement of parents in the reading process, particularly as characters interacting with VAs during read-alouds,

may also yield benefits akin to those derived from reading practices combined with theatrical performance, fostering children's positive attitudes toward reading and enhancing their reading skills (Mastrothanasis et al., 2023).

The purpose of this study was to gain an understanding of parent-child-VA reading as a new mode of joint reading, focusing on both the experiences and values from the perspectives of children and parents. Specifically, the study aimed to address the following research question:

- What are children's reading experiences in the context of parent-child-VA joint reading? (Section 5.1)
- What are parents' perceived values of parent-child-VA joint reading? (Section 5.2)

To answer these questions, we conducted a usability study involving 11 pairs of parents and children (N=22), focusing on children between the ages of 3 and 6 years. We focused on this age group because it is a critical period for establishing vocabulary and literacy skills, and since children in this range are not yet able to read independently, joint reading with adults is essential for supporting language and cognitive development (Griffin et al., 1998; Wijns et al., 2023; Dore et al., 2018). We asked the participating parent-child pairs to read e-books in three different scenarios: 1) a parent as a narrator with two VAs, 2) a parent as a character with two VAs, and 3) a parent and a child, each as a character with one VA. Through this user study, which enabled parents to explore different roles — a narrator and a character, we aimed to gain a comprehensive understanding of how joint reading with VAs might benefit the experiences of both parents and children, depending on the parents' roles in the story: either as a voice actor who needs to dramatize and act out a character, or as a storyteller who leads the narrative.

In addition, through interviews with parents, we evaluated whether this additional task, on top of reading some of the lines, was challenging for parents compared to traditional joint reading with printed books, where they typically read the entire book on their own. Unlike narrated e-books or audiobooks, TaleMate adds the task of having parents take turns with the VAs and pace their reading by clicking through the book to trigger the VAs to read aloud at relevant points. Overall, we aimed to understand how parents perceived the different role-playing scenarios and the additional pacing task.

Lastly, we asked the children questions directly to evaluate their perceptions of TaleMate, the stories they heard through TaleMate and their parents, and the voice assignment process. Children's behavior during reading was observed to assess their engagement in joint reading, and they were asked to answer questions to measure their recall of the book's content.

The study results suggest that TaleMate could serve as a useful platform to engage parents in reading e-books with their young children through the integration of multiple voices and the participatory VA selection process. Moreover, we found that children were actively engaged and enjoyed their reading experiences using TaleMate with their parents. Despite some reservations, parents acknowledged the benefits and potential of using the TaleMate system, particularly for emergent readers. The findings of the current study offer specific guidelines for designers and developers interested in enhancing parent-child interactions through e-books.

Our work contributes to understanding the effects of parental involvement in joint e-book reading through read-aloud with VAs and generates design considerations for e-books that involve VAs, parents, and children. We believe that involving parents and children as active participants in the emerging form of AI-based media during read-aloud is a promising way to make human-AI interactions *child-centered*. We expect parent-child-VA reading to create a more en-

gaging and interactive reading environment compared to solely listening to audiobooks or fully relying on AI agents to read e-books.

2. Related Work

This section presents a review of the existing research and technologies focused on supporting early reading activities among young children. We first summarize the benefits and challenges of parent-child joint reading with print and digital books and then provide an overview of intelligent systems designed to encourage children's reading activities.

2.1. Print and Digital Media for Parent-Child Joint Reading

According to Vygotsky's sociocultural theory, social interaction between a child and a more knowledgeable person, such as a parent or caregiver, is considered fundamental for the child to acquire new information and make developmental progress internally by engaging in question-asking and answering processes (Vygotsky and Cole, 1978). Parent-child joint reading involves a social experience where a parent and child focus on the same text and content together (Bus, 2001). Joint reading, sometimes used interchangeably with joint attention, has been highlighted by many developmental theorists and researchers as a key aspect involved in language development, including early vocabulary and literacy, as well as cognitive and social development (Moore and Dunham, 1995; Ninio and Bruner, 1978).

Specifically related to media, parent mediation theory provides a framework for examining how parents influence and guide children's media use, including digital and print media (Valkenburg et al., 1999; Livingstone and Helsper, 2008). Historically, parent mediation theory has primarily focused on parents' roles in mitigating the negative effects of media use (e.g., social media, tablets, intelligent agents), especially for older children and adolescents (Clark, 2011). Parent-child joint reading and its implications for young children's learning—rather than focusing solely on negative media effects—have often existed separately from the literature on parental mediation theory. For example, parents can use an active mediation strategy to manage, interpret, and engage with media content alongside their children. This includes active mediation, co-viewing/co-use, and restrictive mediation, emphasizing the interaction between parent and child around various forms of media use and highlighting the importance of active parental participation (Connell et al., 2015).

The experience of reading an e-book with a parent falls under the broader concept of *joint media engagement*, where parents and children share experiences with digital learning resources (Walsh, 2017). For example, when parents read an e-book aloud to children or expand on the content by asking questions or engaging in discussions with children about the book, these activities are all considered part of the joint e-book reading experience between parents and children. Empirical research indicates that joint reading, involving parent-child or educator-child dynamics, plays a crucial role in children's language and cognitive development (Krcmar and Cingel, 2014; Torr, 2020). During parent-child joint reading, children are frequently introduced to language slightly above their current level, helping them expand their vocabulary, improve comprehension of various subjects, and cultivate critical thinking skills (Ard and Beverly, 2004; Kleeck, 2003; Torr, 2020; Wesseling et al., 2017). Children's reading aloud during joint reading has also been shown to provide similar benefits (Duursma et al., 2008). Such positive effects of joint book reading have been found in diverse families across socioeconomic strata (Bus et al., 1995).

The quality of interactions children experience during joint reading plays a crucial role in their knowledge acquisition (Lennox, 2013), which holds true for both print books and e-books.

However, differences between print books and e-books have been identified in their impact on parent-child interaction and children's comprehension and learning (Furenes et al., 2021). Many e-books feature audio narration, which shares similarities with audiobooks. Audiobooks have been shown to captivate listener attention, enhance understanding, foster pronunciation accuracy, and positively influence reader-listener attitudes (Best and Clark, 2021; Miranda-Cueva and Cabanillas-Carbonell, 2020). Despite these advantages, audiobooks also have inherent limitations for involving parents in reading. The receptive and linear nature of audiobooks may not inherently encourage active participation in joint reading experiences (Marchetti and Valente, 2018). Indeed, e-books with minimal features, such as simple voice-overs, have been shown to be less effective than print books for children's story comprehension and vocabulary learning in a recent meta-analysis study (Furenes et al., 2021). Thus, a multimodal approach combining audiobooks with traditional books has emerged, leading to the development of novel products in commercial sectors, such as VoxBooks (Publishing, 2014) and Playaway (Products, 2016). However, the effectiveness of such an approach requires further examination.

E-books with voice-overs can improve comprehension and vocabulary learning by intentionally integrating auditory cues and multimedia effects that are relevant to the content, suggesting their potential to emulate the benefits of joint reading using traditional print books (Bus et al., 2015; Furenes et al., 2021; Reich et al., 2016). Further, such intentional design in interactive applications has been shown to promote richer conversations between parents and children (Rowe et al., 2021; Griffith et al., 2023; Chang and Breazeal, 2011). Although digital media have introduced innovative opportunities for enhancing joint reading experiences, electronic devices are still rarely used for joint reading compared to traditional print books (Etta, 2019; Nicholas and Paatsch, 2021). Specifically, there remains a crucial need for solutions that more effectively bridge the gap between traditional interaction dynamics and the capabilities of modern technology. This underscores the importance of developing digital reading platforms that not only enrich the interactive experience for children but also incorporate diverse approaches to more meaningfully engage both parents and children in the reading process.

One underexplored aspect of interactive digital media is the theatrical and collaborative component of storytelling. Compared to traditional e-books and audiobooks, which typically offer linear voice narration, converting joint reading into a theatrical performance that parents and children can design has the potential to create a new avenue for children's participation. Many studies have confirmed that readers' theater, in which actors present dramatic narrative readings, positively contributes to the development of students' reading skills (Mastrothanasis et al., 2023). To that end, we advocate for changing the perception of joint reading into a theatrical performance by allowing parents and children to co-design their reading experience, deciding which character to play. This transition will be reflected in our first design goal: **(Design Goal 1) Support Co-Designing a Theatrical Reading Experience**. For example, TaleMate allows parents and children to assign a VA to each character in the story collaboratively or to play one of the characters themselves. Our approach aligns with the active mediation strategy, where parents converse with their children about the content they will consume and discuss the VAs' roles in their joint reading session, giving parents and children an opportunity to form their perceptions of future media that involves intelligent agents (Valkenburg et al., 1999; Nathanson, 1999). We anticipate that the development of identifiable characters will significantly enhance the reader's immersion into the narrative, based on the extended transportation-imagery model (van Laer et al., 2013).

2.2. Exploration of Intelligent Systems in Children's Reading Experience

Efforts to enhance children's reading experiences have led to the development of various intelligent, interactive solutions. Emerging evidence suggests that interactive voice agents hold promise as social partners for children, as children perceive them as trustworthy, and their interactions can be modeled after conversation (Xu, 2023; Marchetti and Valente, 2018; Xu and Warschauer, 2020). In these studies, voice agents (VAs) have been used as conversational partners, rather than as sole readers, to introduce new interactive and conversational elements to the e-book reading experience, addressing some of the inherent limitations of non-interactive e-books.

VAs have also been utilized to augment the interaction between parents and children, rather than simply substituting the parent's role. As one of the early studies in this area, Chang and Breazeal (2011) introduced interaction techniques aimed at fostering storytelling behaviors during parent-child joint reading, though it did not incorporate voice-assisted reading(Chang and Breazeal, 2011). Many of recent studies facilitated parent-child joint reading through AI-initiated questions during reading. Floppy is one such example, a virtual assistant designed to accompany a child reader, listen to their parents read aloud, and then ask questions related to the book's content at specific moments in the narrative (Lin et al., 2022). Another example is Rosita, a system designed for bilingual children whose parents are native Spanish speakers, presenting culturally relevant questions throughout the reading process to empower parents in conversations about the topic (Xu et al., 2023). StoryBuddy is another system that leverages LLMs to provide questions for parent-child discussions while offering parents the option to read the book themselves or use the VA to read it aloud (Zhang et al., 2022). All of these studies share the goal of maintaining the mode of parent-child joint reading while augmenting it with interactive content or intelligent agents that ask questions about the content.

The current literature has focused on AI's ability to facilitate question-asking but has overlooked the opportunity to involve a voice agent as a reading partner alongside parents for a more diverse, expressive, and theatrical performance. Reading practices facilitated by multiple agents — both parents and VAs—have the potential to foster children's positive attitudes toward reading and enhance their reading skills through theatrical and dramatic performance (Mastrothanasis et al., 2023). Although relying solely on VAs for read-aloud activities may alleviate the burden on parents, it diminishes parental involvement in the text of the book. In addition, previous studies have shown that parents desire to engage in read-aloud activities (Storer and Branham, 2019). Our work attempts to address this gap by offering a new approach for parents to interact with e-books during joint reading with their children, incorporating balanced support from VAs. This knowledge gap resulted in one of our design goals: **(Design Goal 2) Facilitate Joint Turn-Taking Reading with Voice Agents.** Therefore, TaleMate aims for a human-AI collaboration model where parents and VAs read aloud in a complementary and cooperative manner.

3. TaleMate: Design of the System

To explore the integration of parents in promoting joint reading in e-book contexts, we developed TaleMate, a e-book with voice narration that involves parents in joint reading experiences by supporting read-aloud practices. We also used it to identify design implications, as well as parental and children's perceptions of this new mode of interaction. As discussed in the related works, we considered the following two design goals for developing TaleMate:

- Design Goal 1: Support co-designing a theatrical reading experience.

Figure 1: Screen facilitating the assignment of roles (4) to book characters (6)

- Design Goal 2: Facilitate joint turn-taking reading with voice agents.

3.1. [DG1] Support Co-Designing a Theatrical Reading Experience

TaleMate offers an enriched reading experience by allowing parents and children to co-design their reading experience. More specifically, they can collaboratively assign a voice agent to one character and themselves to another. For voice assignment, we prepared six distinct voice agents, each with a unique artificial voice. The system provides more voice agents than the number of characters in the books used for the study. This availability of multiple voices addresses a limitation found in existing e-books with voiceovers, where the narration is typically provided by a single voice actor. Using speech synthesis with different genders, tones, and accents helps mitigate the limited expressivity of linear voiceovers and enables richer character portrayals. Assigning distinct voices to each character makes it easier for listeners to distinguish between them. Moreover, involving the parent as an active participant in read-aloud activities within this multi-voiced narrative elevates parental engagement, ensuring a more interactive and immersive reading environment. TaleMate aims to move beyond the receptive use of e-books with voiceovers, offering a dynamic and engaging joint reading experience that is both educationally beneficial and emotionally resonant.

Within the system, the voice agents are called “Mates,” each represented by an animated avatar designed after a book. We use the Google Text-to-Speech API to read digitized book data and provide six unique voices that can be assigned to each character in the e-book. The multi-voice feature draws on research suggesting that the integration of audio cues and multimedia effects can enhance reading comprehension (Bus et al., 2015). The voice assignment page is designed to facilitate active discussions between parents and children as they assign a particular voice to each character in the book. As illustrated in Figure 1, both the parent and child (Figure 1-7) can assign different voices to each character and to themselves—both the parent

and the child—thereby promoting joint reading practices (Parents/VAs or Parents/Children/VAs) with specific roles. This design enables users to shape their reading experience constructively by assigning voices to each character. The avatars representing the parent and child can be personalized using the Avataaars-generator library (Lin, 2021). Users can create their avatars prior to the voice assignment step, which promotes inclusive avatar creation and makes the process interactive for both parent and child. It is also possible to assign voices only to VAs, so the entire book can be read solely by VAs, similar to e-books with multiple voice actors.

Figure 1 presents the screen where users can assign voices in our system. Users can assign voices to themselves and VAs for each character (Figure 1-6) by dragging and dropping each VA icon (Figure 1-4) from the left panel into one of the boxes, thereby assigning the specific role to the respective character. To further assist the parent user, instructions are displayed at the top of the page (Figure 1-2). Users can click the "Play" icon (Figure 1-5) below each voice agent icon to hear the voice of the agent saying, "Hello, I am Mate X," where X denotes the Mate's number. This feature allows users to listen to each Mate's voice before making a selection. Upon completing the voice assignment, users can click the button at the top-right corner (Figure 1-3) to proceed to the story, or click the button at the top-left corner (Figure 1-1) to return to the home screen. The pre-reading activity provided by the TaleMate voice assignment step facilitates a novel interaction between parents and children: previewing the characters that appear in the book in advance, actively thinking about matching voices with characters based on the cues that children and parents have (e.g., names, visuals, etc.), and initiating parent-child discussions on these topics.

3.2. [DG2] Facilitate Joint Turn-Taking Reading with Voice Agents

In TaleMate, parents read aloud content in a manner similar to reading a printed book aloud to their child. However, when reading with VAs, parents will portray only one of the characters, reading aloud that character's lines, while the remaining characters will be read aloud by VAs or children. Having multiple voice actors, including VAs, presents the challenge of coordinating when a parent or child should read and when a VA should read, ensuring a seamless transition between voices. Therefore, the second goal is to offer a user interface design that facilitates joint turn-taking reading with VAs.

In our design, parents can manually pace their storytelling by interacting with the system. If a parent presses the "Next" button (Figure 2-9) or the spacebar on the keyboard, the system will have the assigned voice agent read the corresponding line. If it is the parent's or child's turn, the system will highlight the line in green or orange, respectively, and wait for them to read it before pressing the button or the spacebar again.

The system is designed for users and voice agents to take turns reading dialogues in the book, based on the character assignments made on the character selection screen (Figure 1). For example, in Figure 2, when the user clicks the "Next" button, the parent's character gets the first turn and reads the narrator's dialogue. Once the parent finishes, they can click "Next" again, and it will be "Mate 5"'s turn, who plays the role of Clara. The cycle continues with each click of the "Next" button, giving the respective characters their turn. Depending on when the parent presses the button, they can pace the reading on their own and, therefore, have the freedom to pause for conversation about the book. To facilitate revisiting passages, the system provides various navigation options (shown in Figure 2-8, 2-9, and 2-10).

We did not use automated pacing with existing speech recognition algorithms (e.g., a forced-alignment algorithm (Katsamanis et al., 2011)) to align parents' or children's verbal utterances with the script. The design rationale behind the absence of automatic pacing is that it could

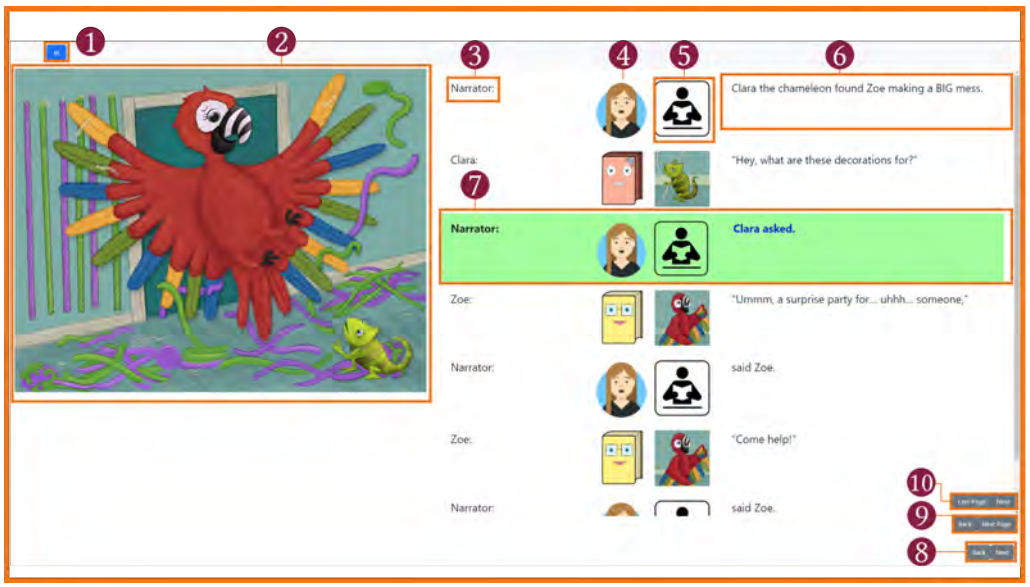


Figure 2: Reading screen where the user and voice agents read the story

discourage natural conversation during reading if voice agents resume reading immediately after parents or children finish their parts. Moreover, failures in speech recognition could compromise the reading experience and potentially frustrate users. We note that it is beyond the scope of this paper to address how to make speech algorithms resilient to recognition errors caused by various factors (e.g., children's interruptions, accents, and domestic environment noise). With manual pacing, parents and children can transition to parent-child interaction whenever they wish, such as through questions, answers, or conversation during reading—an interaction that naturally and frequently occurs in joint reading. In turn, parents will pace the reading by clicking through the e-book manually.

3.3. Design Development Process

The design process for TaleMate was iterative and collaborative. The authors of this paper, comprising experts in human-computer interaction (HCI) and child development, engaged in ongoing discussions to shape the initial design. The first author developed multiple iterations of TaleMate, incorporating feedback from the group to refine and enhance the prototype. Additionally, pilot studies were conducted to uncover and address usability issues, ensuring a more robust and user-friendly system.

The prototype introduced in this section is not currently available, but the second prototype, which was improved based on the study findings, is available for a demo at `https://talemate.cs.vt.edu/`. For more detail about the improvement, see Section 5.3.

4. Study Design

We conducted a user study with parents and children to evaluate TaleMate's effectiveness in enhancing joint reading experiences. Both qualitative and quantitative data were collected

to understand their user experience. All sessions were video recorded to analyze the behavioral patterns of parents and children and to document the exit interviews. Additionally, parents completed a pre-study survey to provide demographic information and details about their media usage.

4.1. Selected Books

For this design, we integrated the book "Birthday Beeps and Boops" from The Pattern Pals series, published by the Purdue Science and Stories Collaborative (Wijns et al., 2023). This book was selected for its age-appropriate content for our target audience (ages 3–6) and its use of distinct dialogues for each character, rather than relying solely on a narrator. Additionally, the three books in the series share a consistent structure featuring two characters—Clara (a chameleon) and Zoe (a parrot)—along with a narrator, making them well-suited for our system. In a previous study, the selected book, enhanced with embedded dialogic reading questions, was shown to be effective in improving young children's patterning abilities (Wijns et al., 2023).

4.2. Procedure

Parent-child pairs participated in two 30-minute Zoom sessions. Prior to the first session, parents completed an online survey covering their demographics, media usage, and reading habits. During the Zoom sessions, parents were instructed to assume different roles (see 3 for an example and detailed description). The two sessions were scheduled at least one full day apart to accommodate children's limited attention spans with shorter sessions of approximately 30 minutes each. This interval also allowed parents time to familiarize themselves with the new technology and simulated repetitive reading—reading the same book multiple times—a common practice among early readers. After each session, we conducted a semi-structured interview to gather participant feedback and reactions.

4.2.1. Pre-study Survey

The online parent survey consisted of four sections, comprising a total of 77 questions. These sections covered informed consent, demographic information, media usage patterns, and preferences for avatar design. In particular, the survey explored how parents engage with their children across various media formats, including traditional books and e-books.

4.2.2. Session 1: Reading as a Story Narrator

After obtaining informed consent, the **Training Task** was conducted to learn how to use TaleMate by using the book "The Pattern Pals: Sleepover Similarities.", which was different from the target book for training purposes. Next, parents started **Task 1 (Parent as a narrator)** with the target book "The Pattern Pals: Birthday Beeps and Boops". The reading task ended with interviews to gather feedback from both parents and children. During the reading task of Session 1, parents were assigned as the "Narrator" of the story, and the VAs that parents and children chose performed the two main characters, Zoe the chameleon and Clara the parrot.

4.2.3. Session 2: Reading as a Main Character With/Without Child

The second session omitted initial training and had parents read the same book but as one of the book's characters by executing **Task 2 (Parent as a main character)**. Parents could choose either of the two main characters available during the task. The final task in Session 2 involved children in the reading process: **Task 3 (Parent and Child as main characters)**. This task

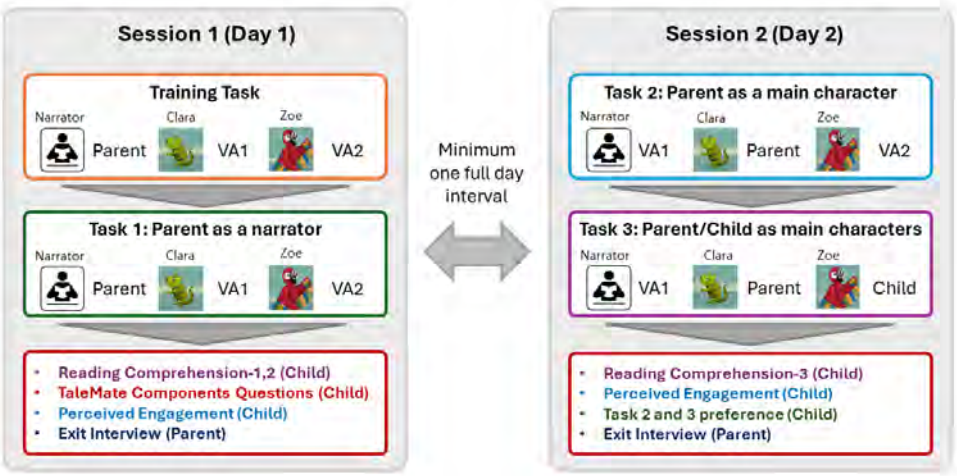


Figure 3: Tasks performed during each session to evaluate TaleMate.

was designed to explore the potential of involving children in theatrical reading. Since the study included children aged 3 to 6 years, parents were asked whether their children could read the text aloud. If children were not yet readers, parents were instructed to read the sentence first and then have their children repeat it. To keep the session brief — approximately 30 minutes, participants read only the first five pages of the book for Task 3. During this session, parents could decide whether to play Clara or Zoe, with the other character role assigned to the child. The Narrator role was performed by a VA of their choice. After completing both tasks, an interview was conducted to examine the effects of role changes and children's involvement in the interactive experience. In both sessions, parents were instructed to behave as they typically would when reading a book with their children. They were also reminded that they could pause and discuss the story, mimicking real-life interactions.

4.3. Participants

The remote study utilized the Zoom platform and involved 11 parent-child pairs. We recruited children between 3 to 6 years old, as our focus was on young children's experiences in joint reading with parents, and this age range aligned with previous studies we were building upon (e.g., (Dore et al., 2018; Wijns et al., 2023)). Recruitment was conducted through university email listservs and local social media posts. The eligibility criteria included having a child in the target age range and a parent who speaks English.

The children had an average age of 5 years (SD = 1.14); 9 out of 11 were girls, and 2 were boys, the ages of all the children can be found in Table 1. All parents were between the ages of 31 and 40 years, except for P7 and P8, who were in the 41-49 age range at the time of the study. All participating parents, except for P5, were female. The study resulted in 22 sessions—two sessions per parent-child pair, with an average duration of 33.25 minutes (SD = 6.86) per session.

Child Code	Age	Gender	Read
C1	6 years and 3 months	Male	Yes
C2	3 years and 4 months	Male	No
C3	5 years and 1 month	Female	No
C4	6 years 1 month	Female	Yes
C5	3 years and 11 months	Female	No
C6	5 years and 5 months	Female	Yes
C7	5 years and 11 months	Female	Yes
C8	5 years and 3 months	Female	Yes
C9	6 years and 7 months	Female	Yes
C10	6 years 10 months	Female	Yes
C11	6 years and 5 months	Female	No

Table 1: Demographic Information of Participating Children

We calculated the children's technology and reading exposures based on the parents' self-reported values. Regarding media use, the 11 child participants spent an average of 2.00 hours watching videos on weekdays and 3.17 hours on weekends. Of these, 8 out of 11 watched videos with parental supervision. Furthermore, children spent an average of 0.25 hours playing video games on weekdays and 0.69 hours on weekends, with 8 out of 11 playing video games under their parents' supervision. In terms of using voice assistants, 7 children spent an average of 0.16 hours on weekdays and 0.35 hours on weekends. Two children (P7, P9) used voice assistants extensively, while the others used them for less than 5 minutes per day, if at all.

Regarding reading habits, parents reported that children spent an average of 0.39 hours reading or listening to e-books at home on weekdays and 0.37 hours on weekends. Only 6 out of 11 children used e-books, and five of them engaged in this activity with their parents. Additionally, all 11 participants read printed books regularly, averaging 1.98 hours per week on weekdays and 1.52 hours on weekends. Furthermore, 10 out of 11 children read printed books with their parents. In addition, we had 4 children who were non-readers and 7 who were early readers.

4.4. Interview and Survey Questions

We evaluated the system by verbally asking children questions about various components of TaleMate. We adopted a two-part questionnaire format to gauge children's impressions of TaleMate. Initial queries (e.g., "Did you like X?") solicited binary "Yes" or "No" answers, with follow-up questions such as "How much did you like X?" offering three choices: a little, some, and a lot. Figure 4 presents the data graphically. This technique was adapted from previous work, which found that the two-part question format is more comprehensible for younger participants (Brink et al., 2019; Severson and Lemm, 2016).

The questions were divided into three overarching categories. The first category was the TaleMate Component Evaluation, which focused on the story, Mates' visuals, and Mates' voices (see Appendix A.2). This category, asked at the end of Session 1, included the **Tutorial Task** and **Task 1 (Parent as a narrator)**. The second category of questions, asked at the end of Session 1 and Session 2, included **Tutorial Task**, **Task 1 (Parent as a narrator)**, **Task 2 (Parent as a main character)**, and **Task 3 (Parent as a main character)**. We also asked the children to compare the modes used in Task 2 (Parent as a main character) and Task 3 (Parent and Child as main characters) with the question: "Which one did you like better, listening to the story (Task 2)

or reading together (Task 3)?". The last category was the Recall questions (see Appendix A.1), which were asked at the end of Session 1 and included the **Tutorial Task** and **Task 1 (Parent as a narrator)**. For all the Recall questions, we showed pictures of the characters along with their names, considering that children might not remember the characters' names but could connect elements of the book with the pictures.

Due to limited statistical power and the absence of randomization, our quantitative data do not provide sufficient support for inferential statistics. However, they complement the qualitative insights by offering additional information on participants' engagement per session.

4.5. Behavioral and Interview Data Analysis

Zoom video recordings were used to examine children's visual attention during the sessions and to analyze parent interviews. Interview data were also transcribed for further analysis. Videos were annotated by independent coders for gaze coding to measure whether participants were paying attention to the screen. We used the duration of children's gaze to the screen as a proxy for children's engagement, based on prior research indicating a strong correlation between gaze duration and user engagement level (Nakano and Ishii, 2010). This duration was calculated as a percentage by summing the time when the children gazed at the screen and then dividing it by the total reading time.

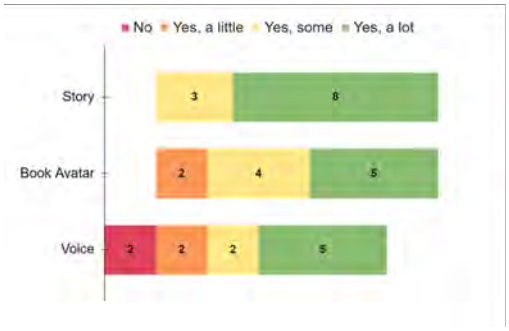
Regarding interview data, the first author transcribed the interviews and conducted a thematic analysis to uncover trends and insights into the user experience (Clarke et al., 2015). The interviews included 22 open-ended questions divided into two sessions, resulting in 242 responses (22 responses per adult). After the initial analysis, the first author identified a set of 48 codes. The first author and the corresponding author then regularly met to refine the codes and identify emerging themes. We also conducted the same analysis procedure for the conversations between parents and children while reading the book during the sessions.

4.6. Ethical Considerations

We designed the study protocol to address any ethical issues that could arise during the research. At the beginning of the study, we explicitly informed participants that their sessions would be recorded and observed, addressing any privacy concerns related to using Zoom and audio/video/screen recording. Participants were also notified that they could withdraw from the study at any time and that their data would be removed from the research project if they chose to do so. Upon notification, we sought parents' consent by providing them ample time to review the consent form. Additionally, we verbally sought the children's assent. When it seemed unclear or if the child was too young, we asked the parents to check if their children would be okay with participating in the study so that they could communicate their intentions directly with their parents rather than answering the study moderator. During the entire process, the study moderator spoke in a child-friendly and positive way (e.g., "We are going to play a game today [when asking questions]."), with a separate introduction to the child to minimize any pressure or confusion on the participating children. The recording was started after we obtained their consent and assent. The entire protocol was reviewed by the Internal Review Board (IRB) of the author's university.

The storybooks used in this study were originally developed and published by the Purdue Science and Stories Collaborative (Wijns et al., 2023). We obtained permission to use the materials prior to the study. The system is available in an open-source repository at `https://github.com/echo-lab/CA`.

(A) Children's evaluation toward TaleMate components



(B) Children's engagement across the two different sessions

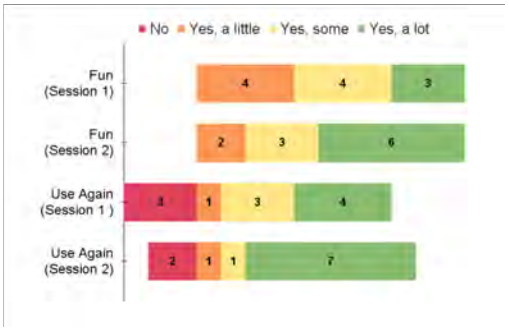


Figure 4: (A) Results obtained from interviewing children about the components of systems (Session 1: Parent-VA reading) and (B) about the engagement metrics in using the systems for Parent-VA reading(Session 1: Training Task and Task 1) and Parent-Child-VA reading (Session 2: Tasks 2 and 3)

5. Results

*5.1. **RQ1: What are children's reading experiences in the context of parent-child-VA joint reading?***

In this section, we outline results based on how TaleMate affected children's reading experience in various aspects: their preference, engagement, and interaction with parents.

5.1.1. Children's Attitudes Towards TaleMate

All 11 children responded to the interview questions in Session 1, which asked about their preference for each component in TaleMate: Story, Avatar Visuals, and Voices. All of these 11 children (100%) said they liked the story (Yes, some = 3; Yes, a lot = 8) and about the Mates' visuals (Yes, a little = 2; Yes, some = 4; Yes, a lot = 5). In contrast, 9 out of 11 children liked the voices used (Yes, a little = 2; Yes, some = 2; Yes, a lot = 5), and two children did not like them by saying No to the question (See Figure 4 - (A)). These results indicate that the participating children had positive marks on the system, with a few exceptions (i.e., the two children who responded "No") on the VAs' voices.

The second category focused on how children perceived the system after the reading session. All children responded positively to whether they had fun using TaleMate in both Sessions. Furthermore, most children desired to reuse the platform for reading, except for a few participants. Notably, more children preferred the reading experience in Session 2 than in Session 1 (See Figure 4 - (B)), suggesting that children's experience with VAs may change as they get more familiar with how it works.

By the end of the second session, we asked the children which they preferred: the Parent-VA Reading Task (Task 2) or the Parent-Child-VA Reading Task (Task 3). Eight out of 11 children preferred the mode in which they are involved in reading, two preferred the Parent-VA Reading configuration, and one preferred both equally.

5.1.2. Engagement: Gaze, Recall, and Parents' Assessment

The gaze annotation results indicated that children were engaged in the reading process and that TaleMate maintained children's levels of attention across the activities. The results indicated that, on average, children remained focused on the screen 89.81% of the time (SD=15.30%)

during the sessions. These high levels of visual attention indicate TaleMate's potential to actively engage its young audience. We conducted a two-tailed paired t-test to determine whether there was any significant difference between the tasks or activities. Our analysis revealed no significant difference between tasks (Task 1 and 2) and activities (reading vs. voice assignment).

We also assessed whether they comprehended the story and could recall key information as an indicator of their engagement with the reading. Eight out of 11 child participants correctly answered the first question (Q1: Whose birthday was it?). In addition, all children answered the other two questions correctly as well (Q2: Clara's favorite colors, Q3: the character that their parents played), yielding 91% accurate rate on average. Most children correctly answered all questions, which suggests their engagement level was sufficient to recall some key elements in the story.

We also asked parents if they thought their children were engaged in reading based on their experience across both sessions. According to feedback from 9 parents, they consistently assessed that their children were engaged in both sessions, even if the children were only listening.

> *"I think she was pretty engaged for never doing something like this before because she's not used to it." (P10)*

> *"He seemed more engaged, you know, repeating the lines. He's never done that before." (P1)*

Two parents said their children were disengaged and kept getting distracted in one of two sessions. One parent mentioned that her child had their head down for most of the reading session and was perceived as fidgety. The other indicated their child was a little tired because they had a long day, and the study occurred in the late afternoon.

5.1.3. Parent-Child Conversation

We analyzed the conversation between parents and children to see how parent-child-VA joint reading brought and expanded the interaction that parents and children would have in their usual joint reading without VAs. We observed active conversations between parents and children in the voice-selection process. We found that all 11 children actively engaged in the process of assigning voices during at least one session. During this process, parents and children exchanged dialogues to decide which voices were most suitable for each scenario. For example, in Pair 2, the parent (P2) and child (C2) had the following interaction:

P2: "I'm going to be Zoe or Clara? Which one do you want me to be?"
C2: "Ummm."
P2: "Do you want me to be the lizard or the parrot?"
C2: "I want you to be lizard."
P2: "Okay I'm gonna be Clara. Now we need to assign voices to the narrator and Zoe the parrot. Do you want to listen to voices again?"
C2: [Nodding]
P2: [After playing voices of VAs] "Which one do you want to describe what is going on?"
C2: "The man with the mustache."
P2: "Who do you want to be Zoe the parrot?"
C2: [Pointing to one of the VA images]
P2: "This one?"
C2: "Yeah."

Some children (C7, C8) directly took the lead in assigning voices and controlled the process, while having conversations with their parents. This voice selection process gave children opportunities to learn all the characters that will be in the story and match their names with the character images.

We also observed conversations about the book content during the sessions. We identified six pairs of parents who engaged in conversations with their children while reading the book. In five out of eleven pairs (P1, P3, P5, P8, P11), the parents initiated the conversation by asking their children questions about the book's content. For instance, P8 and C8 had the following exchange:

P8: "What do you think?"
C8: "I think it's her birthday."
P8: "Whose birthday?"
C8: "Zoe."
P8: "You think it's Zoe's birthday? You think Zoe is setting up her own birthday decorations?"
C8: "No."
P8: "So whose birthday do you think it is?"
C8: "Clara?"
P8: "You think it's Clara's? Okay. All right. Let's listen to the story."

In addition, four out of eleven pairs (P1, P3, P8, and P9) had conversations initiated by the children, either by asking questions or describing the book's content. For instance, P3 and C3 had the following interaction:

C3: "Yellow, orange, yellow, orange."
P3: "Do you know what shapes the patterns are? What is that one? Do you remember that one?"
C3: "Yellow."
P3: "That's good, [child name]. But it's a square. And what's this one? Square, circle, square, circle."

However, 3 parent-child pairs did not have any conversation during the reading. During the parental interview, 3 parents said they could have stopped and talked about the story.

> *"I was definitely thinking like, this would be a good time to have a conversation about why Clara's like face was sad and stuff like that." (P2)*

In the meantime, the other new 2 parents told us that they had not had additional conversations beyond reading when they read books in general, which potentially made them behave similarly when reading with TaleMate.

5.2. ***RQ2: What are parents' perceived values of parent-child-VA joint reading?***

In this section, we present parents' perspectives on the value they saw in TaleMate and the improvements they suggested.

5.2.1. Parents' perceived benefits in using TaleMate

Due to the choices that a parent and a child can make together when choosing which character they will play, TaleMate provided diverse modes of reading, which positively influenced their experience.

[A new option for their interactive reading.] From our 11 participants, 8 of them indicated that they were positive about using TaleMate in their reading routine. One of the main reasons why participants were interested in using it was because they perceived it as similar to a printed book in terms of being able to be involved in joint reading, while also valuing its unique and interactive features. For example, P9 observed that since children of the current generation are drawn to electronics, TaleMate could provide a compelling blend of technology and reading, an activity they expect their children to enjoy.

> *"I think she would like it because, you know, kids at this age are all about electronics. And so this kind of brings the electronics, back to the reading." (P9)*

P3 characterized TaleMate as *"an option in your toolbox,"* particularly for children learning to read. Finally, another participant mentioned they might like to use it because it might help them with reading comprehension. Some participants' openness to TaleMate suggests that the platform has the potential to offer children enhanced reading experiences in the near future. In summary, the parents considered that TaleMate offers a delicate balance between the familiarity of traditional book reading and the innovation brought by voice agents, appreciating the added diversity of experiences TaleMate can offer.

[Voice assignment process is fun and easy to use, but can result in too similar voices.] The assignment process offered a new layer of interaction between parents, children, and voice agents. This process has been accepted and praised by parents. Three out of 11 participants found this capability fun for children, while 9 affirmed the ease of use in the voice assignment process.

> *"I definitely like having to assign voices that he was able to pick them out. I think that was fun that he could do that." (P2)*

However, during the assignment process, some parents found that the voices did not match their expectations based on the look of the avatars.

One challenge identified by some of the parents was that some users felt their chosen voices were too similar, making it difficult to differentiate between them. These similar voices may account for the result that the majority of participants (8/10) preferred playing the roles of Clara or Zoe over playing a narrator. They generally felt embodying a character enriched their interactions and gave them "more reasons to elaborate (P3)." Role-playing enabled them to modulate their voices, infuse emotion, and engage more personally with their children, which can be a strength compared to sometimes monotonous, robotic voices expressing emotion. P9 noted that assuming a character made navigating the turn-taking process easier.

> *"I think it was easier with me being a character just because of the voices. I remember the last time I said they were, they [two voice agent's voices] were very similar. So even to me, it was a little confusing who what character was doing what. But if I'm one of the characters, it made it a little easier to know which character was doing what." (P9)*

Therefore, the voice assignment process needed a careful recommendation depending on the number of characters and their interaction so that voices assigned to two characters that interact

with each other could have been purposely distinct. It was easier for the parents to comprehend when they chose two voices with different genders.
[Different voices "make the same story seem different." (P8), but they sounded awkward.] Most parents welcomed the availability of multiple voice options, which allowed for a more personalized storytelling experience. A comment from P10 reflects this sentiment:

> *"It was a good selection because it had different vocal tones for the men and women. So whatever sounded better to somebody, obviously they would click. And I know that's an individual preference, but you have basically a high, middle, and low tone for everybody." (P10)*

While some participants appreciated the diversity of options, the quality and naturalness of these VA voices remain points of contention. Several users expressed dissatisfaction with the quality of synthetic voices, describing them as *"robotic"*, *"awkward"*, *"fake"*, and *"unnatural"*. For instance, P1 and P9 shared their opinions as follows:

> *"I thought they were a little, kind of fake. You know, it was very obviously a computer reading them." (P1)*

> *"I personally found them to be sounding very electronic. Not quite as natural as I would have liked." (P9)*

This sentiment reveals the limitations of the Google Text-to-Speech API used in TaleMate, with six out of the eleven surveyed users noting that the voices were not convincingly human-like.
[Involving their child in reading aloud made it engaging.] During the exit interview in Session 2, all parents except one (10/11) stated that their children seemed more engaged when they were involved in the reading (Task 3) compared to the task where parents were only involved (Task 2). These parents described the experience as "enjoyable", "effective", and "easier" for the children to understand. P8 provides a detailed account of their child's engagement levels throughout the study.

> *"I think in the listening part, she was slightly less engaged because she already knew the story. So I guess I think the very first time it was new, so it did catch her attention the second time she listened. I think, like I said, she'd been through it, she'd already heard it. But then the third time, her participation really again grabbed her attention." (P8)*

Nearly all of the parents (10/11) expressed a positive opinion about involving their child in the reading process. For example, one parent found it both challenging and beneficial to have her early-reader child read along and then repeat the sentences. Parent 2 (P2) mentioned that the "highlighting feature when it is their turn to read" helped the child understand when it was their turn. Based on parents' opinions and perceptions, involving children in the reading process, even by repeating what a parent uttered, left a positive impression on parents. The result suggests TaleMate provides a novel interaction in which a child and a parent can be involved in collaborative reading aloud with the aid of VAs, in cases in which there are more than two characters.

5.2.2. Patents' Suggestions for Improving TaleMate

We received various suggestions from participants for enhancing the functions of TaleMate to meet different needs. Participants offered recommendations aimed at supporting early reading, enhancing personalization, and promoting comprehension and discussion.

[Semi-automation for turn-taking with voice agents.] One usability issue that we found was the parent's manual pacing feature, having to press a button to pace VAs to read the lines. Eight participants reported that the pacing task felt "confusing" or "disconnected" due to the turn-taking design.

> *"It was good, but it was a bit confusing, I'd say, because I thought it was my turn and it was not my turn, so it was a bit confusing." (P4)*

> *"I think it was a little disconnected between when they said their sentence." (P5)*

This disconnection was mentioned, especially when the parents acted as a narrator and often when they had to say a short sentence for a dialogue tag or attribution, such as "Said Zoe." or "Clara asked." (See Example lines in Figure 2.), supposedly immediately after the voice agent reads a line, which is a moment when they are supposed to press a button as well. The delay caused by the parent having to press a button made it sound disconnected, or it created the same kind of awkwardness when a parent had to press the button for a VA to read attributing sentences immediately after the main character's line.

In contrast, 6 parent participants voiced their appreciation for this feature's control when we specifically asked if they would prefer an completely automated app to a manually pacing app.

> *"I like having control. And this goes back to our earlier discussion about pausing and asking questions, you know, maybe about what the story is about. So far, I believe if it's automated, I wouldn't have that opportunity to do that." (P8)*

> *"I thought it was fun to let them click along, so it kind of helped them stay engaged because you hear the sound of me pressing the keyboard. They know there's another line." (P10)*

P10 even allowed their child to click through, which was another engagement for a child and potentially beneficial for comprehending the story, knowing where they are.

Although the next button conferred some control to parents, enabling them to adjust the story's flow based on their child's needs, parents felt that the turn-taking aspect lacked the smoothness of traditional reading. Consequently, some (5/11) suggested adopting a semi-automatic approach.

> *"I would say maybe more automatic. I mean, if you have an adult reader, they can probably keep up with it as long as they're paying attention... I would say if there was like a play button and then it just highlighted when it was my turn, instead of me having to manually like move each sentence along." (P3)*

Similarly, P1 proposed adding a "pause and play" button to minimize interactions, allowing breaks only when users request it, which makes it close to an audiovisual e-book that is similar to a video, except there are lines that parents are asked to read.

[Features for early reading support.] Participants viewed TaleMate as a beneficial tool for young readers and suggested the following functionalities to enhance its educational value.

Two participants believed synchronized word-level highlighting could be invaluable for children learning to read, as they can see their words being highlighted as the voice agents read them.

> *"For slightly older kids that are learning to read, it might be nice to like highlight each word when it's their turn to read so that they see the highlighted word versus just the whole thing highlighted." (P5)*

Along the same lines, P1 envisioned a dynamic system that moves the text forward as children correctly pronounce each word. To assist readers who may struggle with new words, three participants recommend a feature that allows children to click on words to hear their pronunciation, thus aiding vocabulary building.

[Emphasizing personalization.] To make the experience more engaging and relatable, some participants suggested incorporating the child's name into the selection to increase engagement. This would enable the "Mates" to say "Hello *[name of the child]*" instead of the current phrase. Another participant argued for the inclusion of more illustrative materials, as younger children are often more visually oriented. Lastly, P6 shared that it might be beneficial to incorporate the parents' voices into the book, apart from the voices already available.

[Enhancing reading comprehension and discussion.] Participants believed that existing features could be repurposed or enhanced to stimulate deeper engagement with the text. Three participants saw value in leveraging the system's pause feature to prompt questions about the narrative, thereby promoting comprehension. They suggested voice agents should encourage conversation or discussion about the book.

> *"If reading comprehension is one of the goals of this, um, this technology, I wonder if pausing at certain points in the text to ask questions would be beneficial rather than waiting till, you know, to the very end." (P8)*

5.3. Prototype Improvement

Based on the results, we improved TaleMate, taking into account the feedback we received. The second iteration is available at `https://talemate.cs.vt.edu/` The following summarizes the changes made based on the study findings, which are also shown in figure 5.

- We employed semi-automated pacing by replacing the "Next" buttons with a button that toggles between "play," allowing VAs to finish reading until a parent or child's turn. We also simplified the buttons and changed the function of one button depending on the state (shown in Figure 5-11, 5-12, 5-13).
- We applied Speech Synthesis Markup Language (SSML) to the text-to-speech engine to improve the VAs' verbal expressions.
- We added questions from The Pattern Pals books, on all pages (Figure 5-10) for parents to facilitate natural conversation with children. Instead of VAs initiating the questions, we left them open-ended for parents to use (Figure 5-9). Examples can be found in Appendix Appendix B
- When previewing the VA's voice, it will say "Hello, [child's name entered], I am Mate X." instead of "Hi, I am Mate X."
- Highlighting any text in the book will make a VA read the selected text aloud to a child.

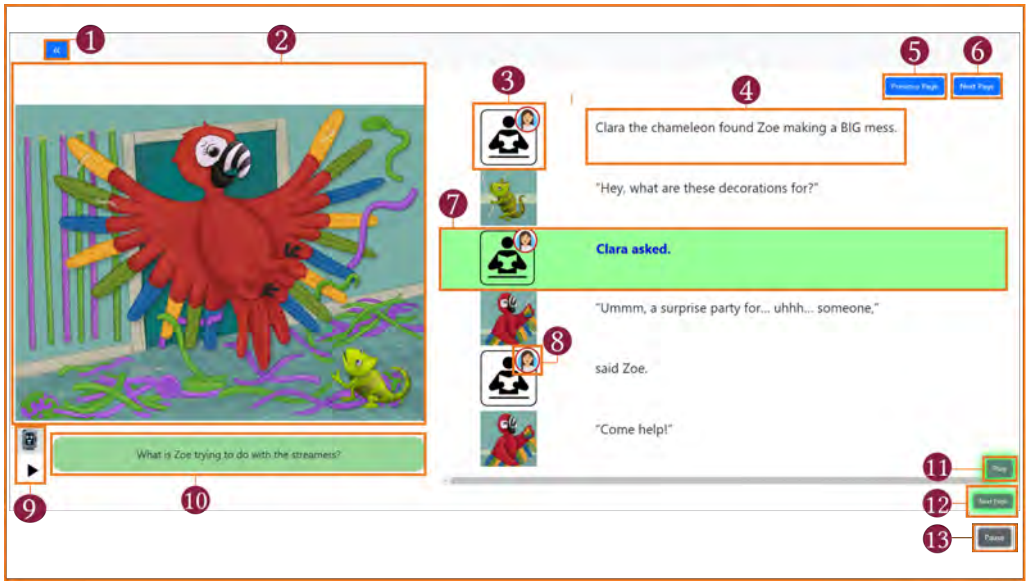


Figure 5: Reading screen where the user and voice agents read the story with the changes incorporated after the user study

6. Discussion

This study delves into how to engage parents in joint e-book reading with voice agents. TaleMate offers parents and children a novel way to actively participate in joint reading by embodying characters through voice agents, enhancing the reading experience, and promoting parent-child interaction. This approach explores the potential for future e-books to foster adult and human-AI collaborative systems and offers design implications for enhancing joint reading experiences in this context.

6.1. Potential of Voice Agents for Engaging Parent-Child Joint Reading

Visual attention is an important component of children's engagement in learning materials (Nakano and Ishii, 2010; Roskos et al., 2012). Further, visual attention while using digital media devices plays a crucial role for reading and learning. That is, well-designed media support learners's visual engagement, helping them to spend more time looking at relevant information, thereby facilitating opportunities for learning and engagement (Kirkorian et al., 2017; Lauricella et al., 2022; Xu et al., 2022). This foundational understanding supports our observation with TaleMate, where we noted significant visual engagement among child participants, evidenced by their gaze duration during the tasks. It is noteworthy that the visual engagement levels observed in our study mirror those reported in a larger-scale study on audiovisual e-books (Richter and Courage, 2017), where the average gaze duration constituted 92% of the entire session. This percentage serves as a comparative benchmark, illustrating that engagement with TaleMate is comparable with broader trends in digital learning tools.

In addition, our preliminary results suggest that most children were able to recall elements from the book after using TaleMate. Although this study did not conduct a direct comparison with

print books, these preliminary results are encouraging, suggesting that it would be worthwhile to measure reading comprehension and systematically compare the differences between virtual and physical modalities (Danaei et al., 2020; Richter and Courage, 2017; Roskos et al., 2012). Parents found reading with TaleMate to be similar to traditional book reading in terms of narrative and engagement but with positive differences, such as the extra interactive and collaborative experience coming from voice assignment and diverse modes of reading with the diverse set of voices. Our work extends existing literature, which has demonstrated how involving parents in joint-media reading can be beneficial for children across media (Duursma et al., 2008; Furenes et al., 2021; Kleeck, 2003; Krcmar and Cingel, 2014; Noble et al., 2019; Sim and Berthelsen, 2014; Torr, 2020; Wesseling et al., 2017), now by involving conversational voice agents and providing more avenues for parents to interact with their children. The results also build upon the recently emerging topic of integrating intelligent systems as reading assistants, as highlighted in various studies (Xu et al., 2023; Xu and Warschauer, 2020; Zhang et al., 2022). In previous studies, voice agents primarily played a question-asking role, whereas in our design, they are incorporated as collaborative partners during read-aloud practices to enhance the joint reading experience.

Furthermore, most children actively participated in the voice selection process and many children engaged in discussions about the book's content. These findings suggest the contribution of parent-child-voice agent (VA) interactions to the collaborative reading dynamic. Conversations were initiated by both parents and children, further reflecting a collaborative interaction. However, in terms of discussing book content, a few pairs did not engage in additional conversations beyond reading, with some parents indicating that this reflected their usual reading practices. These findings suggest that the potential of VAs to facilitate parent-child engagement during joint reading, although the level of dialogue may vary based on individual family reading practices, may require additional support, such as the inclusion of question prompts, which we addressed in our prototype improvement step mentioned in Section 5.3.

In the study, we have witnessed multiple pieces of evidence that TaleMate allowed children to take more agency in the joint reading process: assigning voices to a character, participating in turn-taking reading, and even being able to pace through the story. Even when some of the children were simply repeating what their parents said verbally, the children were highly engaged in Task 3 where they were reading the book with parents and VAs. As emphasized in Bergin (2001), the affectionate quality of the parent-child relationship during the reading may play a more significant role, than the frequency of reading itself, in children's engagement and learning during the reading. While some children were not actively engaged in discussions with parents regarding the story, such elements of our TaleMate system may be able to serve as facilitators to create a more engagement environment, laying an important foundation for supporting children's reading comprehension and word learning. Also, the parents offered many ideas on how we can improve the system for their children's language and cognitive development.

6.2. Designing Considerations regarding Joint Reading Experiences with Voice Agents

Our research objectives aimed to understand how to engage parents in e-books through read alouds to enrich parent-child reading experiences. Our findings augment the existing body of knowledge and offer practical insights into integrating parents and voice agents into joint reading activities.

6.2.1. Modify VAs that Suit the Content of the Book and Vice Versa

Our results indicate a desire for more natural-sounding, "human-like" voices. Previous research indicates the importance of selecting voices that align with a character's personality in storytelling (Cambre et al., 2020). A sentiment analysis of each character could aid in this voice-matching process. For example, if Clara the Chameleon is portrayed as melancholic for much of the story, a voice agent voice could be suggested to reflect this emotion.

Participants also appreciated the diversity of voice options we provided. However, we recommend extending the range and making the selection processes more dynamic. By dynamic, we mean that if a user selects a "low female" tone for one character, the other characters can be recommended with distinct tones to avoid confusion and enrich the storytelling experience.

On the other hand, we note that the content of the book can be adjusted to better align with the theatrical and dynamic nature of the joint reading with VAs. For instance, sentences attributing dialogue (e.g., "Said Clara") can be omitted, as they may cause confusion and are unnecessary given the visual cues provided by character icons and additional prompts indicating who is speaking (e.g., highlighting Clara's picture when the app detects sound from the microphone). Thus, while traditionally essential in book settings to clarify who said what, such sentences can be considered as dispensable with visual, animated, and interactive features.

6.2.2. Guide Parents and Children in Choosing Characters for a Richer and Participatory Reading Experience.

The voice assignment step in TaleMate offered a novel, participatory space for parents and children to discuss and make a connection to the story prior to the reading. Reading with VAs in a multi-character story can take advantage of this multi-agent selection process. Future designers and researchers can also consider improving their reading experiences by guiding the choice of characters that parents and children can opt to read. In our study, the narrator's role often consisted of less rich content like "she said." or "said Clara." According to parents, being a character allows for greater emotional expression compared to being the narrator. Depending on the parents' comfort levels in impersonating a character, the system can encourage parents to read a particular character. Additionally, parents might also benefit from knowing the complexity of each character's lines so they can suggest a particular character to their children. For example, Zoe's lines contained more multi-syllable words than Clara's, making Zoe a better fit for advanced readers or parents, whereas Clara might be more suitable for early or non-readers.

6.2.3. Make Pacing as Easy as Possible and Create Space for Conversation

Participants primarily expressed concerns over the pacing of the shared reading experience, finding it less smooth than reading without voice agents. This was particularly due to the necessity of interacting with the system to advance the text line by line. To address this, we recommend adopting a semi-automatic approach where the voice agent continues reading until it is a user's turn. Introducing pauses for users in the middle of reading could create opportunities for dialogue. Along the same lines, we suggest incorporating mechanisms similar to "Floopy" (Lin et al., 2022) and "Rosita" (Xu et al., 2023), which capitalize on these pauses to ask questions and promote conversation.

6.2.4. Take Advantage of Conversational Interaction for Learning, Language Acquisition, and Beyond

Another avenue for improvement lies in leveraging the system for vocabulary and literacy development. While we did not explore this particular potential, the parents viewed TaleMate

as a potential tool for supporting emergent readers, similar to read-aloud and shared reading. The system can provide additional educational benefits by adding features such as highlighting word by word when voice agents are reading, and allowing children to click on words if they encounter difficulties. While such elements were suggested in e-books, e.g., adding a dictionary, story-related enhancements, and highlighting text (Furenes et al., 2021), the inclusion of VAs affords a conversational interaction for these features that can add intelligence responses in these enhancements. For example, a child can ask different questions while pressing a word on the screen: "Can you read this word again?", "What does this mean?", and "I have a blanket like this, too.", which will lead to different responses from VA, as opposed to a single type of response (e.g., reading aloud the word again).

Furthermore, such interactivity creates room for VAs to facilitate new conversations through diverse types of questions and prompts. For example, a VA could ask a contextually relevant question tailored to the child's immediate environment or activity, such as, "Do you like your blanket?" This type of question can often simulate meaningful conversations between parents and children. Expanding on this by incorporating other types of questions, such as open-ended, descriptive, and causal questions, can further enhance various aspects of children's language learning, including vocabulary growth and oral narrative skills (Pillinger and Vardy, 2022; Hargrave and Sénéchal, 2000)

In conclusion, our research highlights several avenues for the effective incorporation of parents into e-book reading experiences by using voice agents to read part of the story. There are numerous opportunities for innovation, from improving the pacing and fostering dialogue to enhancing educational benefits and creating a more immersive experience. We hope that future researchers and designers can use these insights to develop systems that not only entertain but also educate and connect families.

6.3. Limitations

While our study offers valuable insights into the potential of conversational technologies like TaleMate to enhance parent-child reading experiences, it is important to acknowledge several limitations that could impact the interpretation and generalizability of our findings.

Firstly, the study was conducted via Zoom, with participants sharing their screens and videos. Being aware that they were recorded, participants might have behaved differently than they would when reading alone with their children. A future study could attempt to emulate a more natural reading environment to assess behaviors more accurately.

Another limitation lies in the study's limited ecological validity of the reading environment. The study environment would have not be close to the natural environment where parents and their children read books together: time, place, watching a laptop screen, and their awareness of a study moderator observing them. For example, as stated in 5.1.3, we noticed 5 out of 11 parent-child pairs did not have any conversation about the content. This lack of interaction can be attributed to the Hawthorne effect (Lazar et al., 2017, p.39), wherein individuals might alter their behaviors when they are aware of being observed. In addition, as we conducted only twice with one primary reading task each time, the study may not fully capture how parent-child interactions would have emerged as they become more accustomed to the system. This limitation calls for a future study that allows the deployment of the system in naturalistic environments, such as homes, in a longer term. Such studies can provide insights into how increased familiarity with the app impacts Parent-Child-VA interactions over time. For example, we did not observe children attempting to converse with VAs in this study. A previous study found that children's between the ages of 3 and 10 asked different types of questions to Amazon Alexa, a type of voice

assistant, as they became more familiar with it (Oranç and Ruggeri, 2021). Utilizing dialogic reading questions within the joint-reading sessions with parents using TaleMate could provide even greater learning benefits for children as a result of increased verbal exchanges initiated or actively participated in by children.

Additionally, our result does not provide quantitative evidence that a particular reading mode is more effective than another in terms of parent-child engagement or conversation. The observed increase in positive feedback from children between Session 1 and Session 2 could be attributed to various factors such as a learning effect. Similar considerations apply to parents who preferred playing a character over being the narrator. Randomizing the order of these roles in subsequent larger-scale studies will be needed to provide sufficient quantitative evidence.

7. Conclusions

In conclusion, our study revisits the critical discussion on the evolving landscape of joint reading in the digital age and how this may lead to disconnection between adults and children. We assessed the TaleMate system, a reading platform designed to mitigate the reduced parent-child interactions involving digital reading formats, particularly e-books. Our investigation into the dynamic interplay between parents, children, and voice agents within the TaleMate system has yielded important insights. We found that TaleMate not only captivates children's attention and interest but also promotes parental engagement in the reading process, which has significant implications for addressing parental challenges involved in shared reading with e-books. The system's flexibility, evidenced through personalized settings and voice options, is appealing to a broad spectrum of readers, enhancing its utility.

While identifying areas for refinement in user interaction, the positive reception from parents underscores TaleMate's potential as an integral component of daily reading practices. This potential speaks directly to our initial concerns about the erosion of traditional reading bonds in the face of digital alternatives. By demonstrating how TaleMate bridges the divide between print and digital reading experiences, our study contributes to a deeper understanding of how innovative technologies can support and enhance the pedagogical and developmental benefits of joint reading.

Furthermore, our findings extend the discourse on e-book formats, aligning with and expanding upon the existing literature on digital reading's impact. The implications for the design and application of such systems are relevant, offering a guide for future researchers aimed at enriching the joint reading landscape. In essence, TaleMate exemplifies the transformative power of integrating voice agents in promoting engaging and educational reading interactions, exploring a new era of parent-child reading.

8. Acknowledgements

We thank Dr. David Purpura at Purdue University for generously providing the storybooks used in this study. We also thank the children and parents who participated in the study, as well as the undergraduate students in the Cognitive Developmental Science (CoDeS) Lab for their assistance with transcription and coding.

References

M. Zivan, T. Horowitz-Kraus, Parent–child joint reading is related to an increased fixation time on print during storytelling among preschool children, Brain and Cognition 143 (2020) 105596. URL: https://www.sciencedirect.com/science/article/pii/S0278262620301718. doi:https://doi.org/10.1016/j.bandc.2020.105596.

E. D. Kunkun Zhang, J. Torr, Reading aloud as performance and its representation on television programmes for children, Social Semiotics 26 (2016) 424–444. URL: https://doi.org/10.1080/10350330.2016.1189734. doi:10.1080/10350330.2016.1189734. arXiv:https://doi.org/10.1080/10350330.2016.1189734.

S. Lennox, Interactive read-alouds—an avenue for enhancing children's language for thinking and understanding: A review of recent research, Early Childhood Education Journal 41 (2013) 381–389.

A. C. Schmidt, M. Pierce-Rivera, L. Van Huisstede, S. C. Marley, K. A. Bernstein, J. Millinger, M. F. Kelley, M. A. Restrepo, What's the Story with Storytime?: An Examination of Preschool Teachers' Drama-Based and Shared Reading Practices During Picturebook Read-Aloud, Early Childhood Education Journal (2023). URL: https://link.springer.com/10.1007/s10643-023-01554-z. doi:10.1007/s10643-023-01554-z.

A. G. Bus, M. H. Van Ijzendoorn, A. D. Pellegrini, Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy, Review of educational research 65 (1995) 1–21.

A. Miranda-Cueva, M. Cabanillas-Carbonell, Benefits of using an audiobook application as an educational entertainment tool for children: a review of the scientific literature in the years 2006-2019, in: 2020 IEEE Congreso Bienal de Argentina (ARGENCON), IEEE, Argentina, 2020, pp. 1–4. doi:10.1109/ARGENCON49523.2020.9505385.

M. Krcmar, D. P. Cingel, Parent–child joint reading in traditional and electronic formats, Media Psychology 17 (2014) 262–281. URL: https://doi.org/10.1080/15213269.2013.840243. doi:10.1080/15213269.2013.840243.

P. B. Wesseling, C. A. Christmann, T. Lachmann, Shared book reading promotes not only language development, but also grapheme awareness in german kindergarten children, Frontiers in Psychology 8 (2017) 364.

L. C. Larson, E-books and audiobooks: Extending the digital reading experience, The Reading Teacher 69 (2015) 169–177.

C. A. Ewin, A. E. Reupert, L. A. McLean, C. J. Ewin, The impact of joint media engagement on parent–child interactions: A systematic review, Human Behavior and Emerging Technologies 3 (2021) 230–254.

N. Kucirkova, V. Grøver, The Importance of Embodiment and Agency in Parents' Positive Attitudes Towards Shared Reading with Their Children, Early Childhood Education Journal 52 (2024) 231–231. URL: https://doi.org/10.1007/s10643-022-01434-y. doi:10.1007/s10643-022-01434-y.

M. Nicholas, L. Paatsch, Mothers' views on shared reading with their two-year olds using printed and electronic texts: Purpose, confidence and practice, Journal of Early Childhood Literacy 21 (2021) 3–26. URL: https://doi.org/10.1177/1468798418792614. doi:10.1177/1468798418792614. arXiv:https://doi.org/10.1177/1468798418792614.

R. A. Etta, Parent preferences: E-books versus print books, Reading in the Digital Age: Young Children's Experiences with E-books: International Studies with E-books in Diverse Contexts (2019) 89–101.

M. I. Furenes, N. Kucirkova, A. G. Bus, A comparison of children's reading on paper versus screen: A meta-analysis, Review of Educational Research 91 (2021) 483–517. URL: https://doi.org/10.3102/0034654321998074. doi:10.3102/0034654321998074. arXiv:https://doi.org/10.3102/0034654321998074.

E. Marchetti, A. Valente, Interactivity and multimodality in language learning: the untapped potential of audiobooks, Universal Access in the Information Society 17 (2018) 257–274.

Y. Xu, M. Warschauer, Exploring young children's engagement in joint reading with a conversational agent, in: Proceedings of the interaction design and children conference, IDC '20, Association for Computing Machinery, New York, NY, USA, 2020, p. 216–228. URL: https://doi.org/10.1145/3392063.3394417. doi:10.1145/3392063.3394417.

G. C. Lin, I. Schoenfeld, M. Thompson, Y. Xia, C. Uz-Bilgin, K. Leech, "what color are the fish's scales?" exploring parents' and children's natural interactions with a child-friendly virtual agent during storybook reading, in: Interaction Design and Children, IDC '22, Association for Computing Machinery, New York, NY, USA, 2022, p. 185–195. URL: https://doi.org/10.1145/3501712.3529734. doi:10.1145/3501712.3529734.

Z. Zhang, Y. Xu, Y. Wang, B. Yao, D. Ritchie, T. Wu, M. Yu, D. Wang, T. J.-J. Li, 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, CHI '22, Association for Computing Machinery, New York, NY, USA, 2022, pp. 1–21. URL: https://doi.org/10.1145/3491102.3517479. doi:10.1145/3491102.3517479.

Y. Xu, K. He, V. Vigil, S. Ojeda-Ramirez, X. Liu, J. Levine, K. Cervera, M. Warschauer, "rosita reads with my family": Developing a bilingual conversational agent to support parent-child shared reading, in: Proceedings of the 22nd Annual ACM Interaction Design and Children Conference, IDC '23, Association for Comput-

ing Machinery, New York, NY, USA, 2023, p. 160–172. URL: https://doi.org/10.1145/3585088.3589354. doi:10.1145/3585088.3589354.

K. Mastrothanasis, M. Kladaki, A. Andreou, A systematic review and meta-analysis of the readers' theatre impact on the development of reading skills, International Journal of Educational Research Open 4 (2023) 100243. URL: https://www.sciencedirect.com/science/article/pii/S2666374023000183. doi:https://doi.org/10.1016/j.ijedro.2023.100243.

P. Griffin, M. S. Burns, C. E. Snow, Preventing reading difficulties in young children, National Academies Press, 1998.

N. Wijns, D. Purpura, J. Torbeyns, Stimulating preschoolers' repeating patterning ability by means of dialogic picture book reading., Journal of Educational Psychology 115 (2023) 732–746. doi:10.1037/edu0000756, place: US Publisher: American Psychological Association.

R. A. Dore, B. Hassinger-Das, N. Brezack, T. L. Valladares, A. Paller, L. Vu, R. M. Golinkoff, K. Hirsh-Pasek, The parent advantage in fostering children's e-book comprehension, Early Childhood Research Quarterly 44 (2018) 24–33. URL: https://www.sciencedirect.com/science/article/pii/S0885200618300073. doi:https://doi.org/10.1016/j.ecresq.2018.02.002.

L. S. Vygotsky, M. Cole, Mind in society: Development of higher psychological processes, Harvard university press, 1978.

A. G. Bus, Joint caregiver-child storybook reading: A route to literacy development, Handbook of early literacy research 1 (2001) 179–191.

C. Moore, P. J. Dunham, Joint attention : its origins and role in development, Lawrence Erlbaum Associates, Hillsdale, N.J, 1995.

A. Ninio, J. Bruner, The achievement and antecedents of labelling, Journal of Child Language 5 (1978) 1–15. doi:10.1017/S0305000900001896.

P. M. Valkenburg, M. Krcmar, A. L. Peeters, N. M. Marseille, Developing a scale to assess three styles of television mediation:"instructive mediation,""restrictive mediation," and "social coviewing", Journal of broadcasting & electronic media 43 (1999) 52–66.

S. Livingstone, E. J. Helsper, Parental mediation of children's internet use, Journal of broadcasting & electronic media 52 (2008) 581–599.

L. S. Clark, Parental mediation theory for the digital age, Communication theory 21 (2011) 323–343.

S. L. Connell, A. R. Lauricella, E. Wartella, Parental co-use of media technology with their young children in the usa, Journal of children and media 9 (2015) 5–21.

E. Walsh, The SAGE encyclopedia of out-of-school learning, Sage Publications, Indiana University, USA, 2017. doi:https://doi.org/10.4135/9781483385198.

J. Torr, How 'shared' is shared reading: Book-focused infant–educator interactions in long day-care centres, Journal of Early Childhood Literacy 20 (2020) 815–838. URL: https://doi.org/10.1177/1468798418792038. doi:10.1177/1468798418792038.

L. M. Ard, B. L. Beverly, Preschool word learning during joint book reading: Effect of adult questions and comments, Communication disorders quarterly 26 (2004) 17–28.

A. v. Kleeck, Research on Book Sharing: Another Critical Look, in: On Reading Books to Children, Routledge, New York, 2003, pp. 273–320. Num Pages: 48.

E. Duursma, M. Augustyn, B. Zuckerman, Reading aloud to children: the evidence, Archives of disease in childhood 93 (2008) 554–557.

E. Best, C. Clark, The Role of Audiobooks to Engage Reluctant Readers and Underrepresented Children and Young People. A National Literacy Trust Research Report, Technical Report, National Literacy Trust, 2021. URL: https://eric.ed.gov/?id=ED614814, publication Title: National Literacy Trust ERIC Number: ED614814.

V. Publishing, Vox books. books that talk, https://www.libraryideas.com/vox-books, 2014.

P. Products, The only family of tablets designed for circulation., https://playaway.com/launchpad/, 2016.

A. G. Bus, Z. K. Takacs, C. A. Kegel, Affordances and limitations of electronic storybooks for young children's emergent literacy, Developmental Review 35 (2015) 79–97. URL: https://www.sciencedirect.com/science/article/pii/S0273229714000501. doi:https://doi.org/10.1016/j.dr.2014.12.004, special Issue: Living in the "Net" Generation: Multitasking, Learning, and Development.

S. M. Reich, J. C. Yau, M. Warschauer, Tablet-based ebooks for young children: What does the research say?, Journal of Developmental & Behavioral Pediatrics 37 (2016) 585–591.

M. L. Rowe, R. G. Turco, J. H. Blatt, Can interactive apps promote parent-child conversations?, Journal of Applied Developmental Psychology 76 (2021) 101326. URL: https://www.sciencedirect.com/science/article/pii/S0193397321000897. doi:https://doi.org/10.1016/j.appdev.2021.101326.

S. F. Griffith, S. M. Casanova, J. H. Delisle, Back-and-forth conversation during parent–child co-use of an educational app game, Early Child Development and Care 193 (2023) 1007–1021. URL: https://doi.org/10.1080/03004430.2023.2199173. doi:10.1080/03004430.2023.2199173. arXiv:https://doi.org/10.1080/03004430.2023.2199173.

A. Chang, C. Breazeal, Tinkrbook: Shared reading interfaces for storytelling, in: Proceedings of the 10th International Conference on Interaction Design and Children, IDC '11, Association for Computing Machinery, New York, NY, USA, 2011, p. 145–148. URL: https://doi.org/10.1145/1999030.1999047. doi:10.1145/1999030.1999047.

A. I. Nathanson, Identifying and explaining the relationship between parental mediation and children's aggression, Communication research 26 (1999) 124–143.

T. van Laer, K. de Ruyter, L. M. Visconti, M. Wetzels, The Extended Transportation-Imagery Model: A Meta-Analysis of the Antecedents and Consequences of Consumers' Narrative Transportation, Journal of Consumer Research 40 (2013) 797–817. URL: https://doi.org/10.1086/673383. doi:10.1086/673383, _eprint: https://academic.oup.com/jcr/article-pdf/40/5/797/9500316/40-5-797.pdf.

Y. Xu, Talking with machines: Can conversational technologies serve as children's social partners?, Child Development Perspectives 17 (2023) 53–58.

K. M. Storer, S. M. Branham, "that's the way sighted people do it": What blind parents can teach technology designers about co-reading with children, in: DIS '19: Proceedings of the 2019 on Designing Interactive Systems Conference, DIS '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 385–398. URL: https://doi.org/10.1145/3322276.3322374. doi:10.1145/3322276.3322374.

F.-P. Lin, Avataaars generator, https://github.com/fangpenlin/avataaars-generator, 2021.

A. Katsamanis, M. Black, P. G. Georgiou, L. Goldstein, S. Narayanan, Sailalign: Robust long speech-text alignment, in: Proc. of workshop on new tools and methods for very-large scale phonetics research, volume 1, 2011, pp. 0–4.

K. A. Brink, K. Gray, H. M. Wellman, Creepiness creeps in: Uncanny valley feelings are acquired in childhood, Child development 90 (2019) 1202–1214.

R. L. Severson, K. M. Lemm, Kids see human too: Adapting an individual differences measure of anthropomorphism for a child sample, Journal of Cognition and Development 17 (2016) 122–141. URL: https://doi.org/10.1080/15248372.2014.989445. doi:10.1080/15248372.2014.989445. arXiv:https://doi.org/10.1080/15248372.2014.989445.

Y. I. Nakano, R. Ishii, Estimating user's engagement from eye-gaze behaviors in human-agent conversations, in: Proceedings of the 15th International Conference on Intelligent User Interfaces, IUI '10, Association for Computing Machinery, New York, NY, USA, 2010, p. 139–148. URL: https://doi.org/10.1145/1719970.1719990. doi:10.1145/1719970.1719990.

V. Clarke, V. Braun, N. Hayfield, Thematic analysis, Qualitative psychology: A practical guide to research methods 3 (2015) 222–248.

K. Roskos, K. Burstein, B.-K. You, A typology for observing children's engagement with ebooks at preschool., Journal of Interactive Online Learning 11 (2012) 47–66.

H. Kirkorian, T. Pempek, K. Choi, The role of online processing in young children's learning from interactive and noninteractive digital media, Media exposure during infancy and early childhood: The effects of content and context on learning and development (2017) 65–89.

A. R. Lauricella, F. Aladé, M. Russo, A. Strevett, J. Herdzina, Children's visual attention and comprehension from synchronous video book reading, Computers & Education 191 (2022) 104628. URL: https://www.sciencedirect.com/science/article/pii/S0360131522001993. doi:https://doi.org/10.1016/j.compedu.2022.104628.

Y. Xu, J. Aubele, V. Vigil, A. S. Bustamante, Y.-S. Kim, M. Warschauer, Dialogue with a conversational agent promotes children's story comprehension via enhancing engagement, Child Development 93 (2022) e149–e167.

A. Richter, M. L. Courage, Comparing electronic and paper storybooks for preschoolers: Attention, engagement, and recall, Journal of Applied Developmental Psychology 48 (2017) 92–102. URL: https://www.sciencedirect.com/science/article/pii/S0193397317300059. doi:https://doi.org/10.1016/j.appdev.2017.01.002.

D. Danaei, H. R. Jamali, Y. Mansourian, H. Rastegarpour, Comparing reading comprehension between children reading augmented reality and print storybooks, Computers & Education 153 (2020) 103900. URL: https://www.sciencedirect.com/science/article/pii/S0360131520300993. doi:https://doi.org/10.1016/j.compedu.2020.103900.

C. Noble, G. Sala, M. Peter, J. Lingwood, C. Rowland, F. Gobet, J. Pine, The impact of shared book reading on children's language skills: A meta-analysis, Educational Research Review 28 (2019) 100290. URL: https://www.sciencedirect.com/science/article/pii/S1747938X18305116. doi:https://doi.org/10.1016/j.edurev.2019.100290.

S. Sim, D. Berthelsen, Shared book reading by parents with young children: Evidence-based practice, Australasian Journal of Early Childhood 39 (2014) 50–55. URL: https://doi.org/10.1177/183693911403900107. doi:10.1177/183693911403900107. arXiv:https://doi.org/10.1177/183693911403900107.

C. Bergin, The parent-child relationship during beginning reading, Journal of Literacy Research 33 (2001) 681–706. URL: https://doi.org/10.1080/10862960109548129. doi:10.1080/10862960109548129, publisher: SAGE Publications Inc.

J. Cambre, J. Colnago, J. Maddock, J. Tsai, J. Kaye, Choice of voices: A large-scale evaluation of text-to-speech voice quality for long-form content, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20, Association for Computing Machinery, New York, NY, USA, 2020, p. 1–13. URL: https://doi.org/10.1145/3313831.3376789. doi:10.1145/3313831.3376789.
C. Pillinger, E. J. Vardy, The story so far: A systematic review of the dialogic reading literature, Journal of Research in Reading 45 (2022) 533–548.
A. C. Hargrave, M. Sénéchal, A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading, Early Childhood Research Quarterly 15 (2000) 75–90.
J. Lazar, J. H. Feng, H. Hochheiser, Research methods in human-computer interaction, Morgan Kaufmann, 2017.
C. Oranç, A. Ruggeri, "Alexa, let me ask you something different" Children's adaptive information search with voice assistants, Human Behavior and Emerging Technologies 3 (2021) 595–605. URL: https://onlinelibrary.wiley.com/doi/10.1002/hbe2.270. doi:10.1002/hbe2.270.

Appendix A. Questions Used During Children's Interview

Appendix A.1. Recall Questions:

- Whose birthday was it, Zoe or Clara? (Session 1)
- What are Clara's favorite colors, blue and orange or purple and green? (Session 1)
- Do you remember which one your Mom/Dad was playing? Was it Zoe? Or Clara? (Session 2)

Appendix A.2. TaleMate Components Questions:

- Did you like the story we read today? (If yes) How much did you like it?
- Did you like how the "Mates" look? (If yes) How much did you like it?
- Did you like the voices of the "Mates"? (If yes) How much did you like it?

Appendix A.3. Children's Engagement Questions:

- Was it fun to use TaleMate?
- Would you like to read with your parent and friends again?

Appendix A.4. Preference Between Tasks 2 and 3

- At the beginning of the session your parent read the book and you listened carefully. At the end of the session, you and your parent read the book together. Which one did you like better? Reading together or listening carefully?

Appendix B. Questions Added to the New Prototype

Appendix B.1. The Pattern Pals: Birthday Beeps and Boops

- What is Zoe trying to do with the streamers?
- What color plate is missing?
- Which balloon is the "beep," and which is the "boop" in the pattern?
- What color was the cake they ordered?
- How is the pattern on the cake the same as the pattern of the streamers?

Appendix B.2. The Pattern Pals: Sleepover Similarities

- What is something fun you like to do with your friends?
- Where do you think Clara will hide?
- What is the pattern on the curtains?
- What is another pattern you could make using the shapes on the curtains?
- What is a pattern on the pillow Clara gives Zoe?