

A Home Study of Parent-Child Co-Reading with a Bilingual Conversational Agent

Kunlei He
kunlei.he@uci.edu
University of California, Irvine
USA

Santiago Ojeda-Ramirez
sojedara@uci.edu
University of California, Irvine
USA

Julian Levine
jslevine@uci.edu
University of California, Irvine
USA

Ying Xu
xyying@umich.edu
University of Michigan, Ann Arbor
USA

Kelsyann Cervera
cerverak@uci.edu
University of California, Irvine
USA

Mark Warschauer
markw@uci.edu
University of California, Irvine
USA

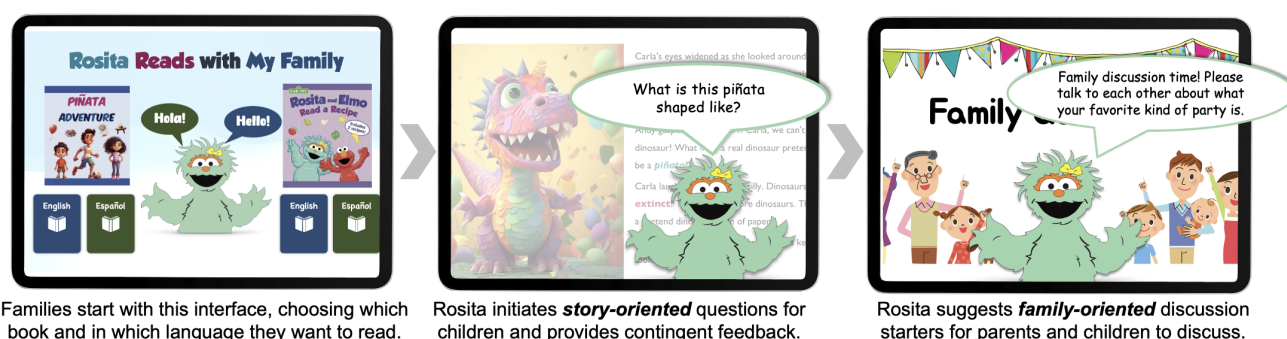


Figure 1: The user interface and flow of the bilingual CA, Rosita

ABSTRACT

Conversational agents (CAs) are increasingly prevalent in children's lives, serving as educational companions, particularly in shared reading activities. While effective for monolingual children's learning, there exists a gap in meeting the unique needs of the rapidly expanding bilingual child population, who face dual challenges of school readiness and heritage language maintenance. Moreover, most current CAs, designed for one-to-one interactions with children, neglect the importance of parents' active participation in shared reading. Our study introduces the development and home deployment of a bilingual CA, integrated within e-books, designed to foster parent-child joint engagement in shared reading, thereby promoting children's bilingual language development. Results of the study indicated high levels of family engagement in co-reading activities over an extended period, with observable language learning gains in children. This study provides valuable design implications for designing effective and engaging CAs for bilingual families.

CCS CONCEPTS

• Social and professional topics → Children; • Human-centered computing → Empirical studies in interaction design.

KEYWORDS

Conversational agents, shared reading, parent-child interactions, bilingualism, home deployment

ACM Reference Format:

Kunlei He, Julian Levine, Kelsyann Cervera, Santiago Ojeda-Ramirez, Ying Xu, and Mark Warschauer. 2024. A Home Study of Parent-Child Co-Reading with a Bilingual Conversational Agent. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3613905.3650836>

1 INTRODUCTION

Young children around the country, regardless of their ethnicity, now have broad access to mobile digital devices with built-in conversational agents (CAs), such as Apple Siri, Amazon Alexa, and Google Assistant. These CAs are capable of understanding and responding to children's speech [22]. Consequently, researchers and children's media experts have been investigating how these consumer technologies could be repurposed for educational benefits. One of the most fruitful lines of research in this area explores the use of CAs to engage children in interactive story reading [5, 26, 31, 38]. Through back-and-forth conversations with CAs,

children demonstrated language learning gains comparable to reading with a human [41]. Despite the promising findings, existing studies predominantly focus on monolingual children [5, 26, 31, 38], overlooking the needs of the growing bilingual child population in the United States. Nearly a quarter of U.S. children live in households where English is not the primary language, with a significant portion from Spanish-speaking families in Latine communities. As of 2020, Latine children made up about 26% of the preschool population, a figure expected to rise. Young Latine bilingual children face unique challenges, as they must balance mastering English for academic success while maintaining their heritage language [35]. This dual-language navigation underscores the necessity for educational tools like bilingual CAs that cater to their specific language development needs.

To address the language development needs of bilingual children, we identified two key areas for bridging this gap with CAs. Firstly, while current CAs used in shared reading have shown effectiveness, they lack bilingual capabilities. Thus, we enhanced these CAs by incorporating bilingual functionalities, enabling them to interact with children in their preferred language and provide tailored feedback, thereby fostering bilingual learning. Secondly, existing CAs are mainly designed for one-on-one interactions with children, which might inadvertently limit parental involvement, despite its proven benefits in enhancing children’s learning [12]. To rectify this, we adapted CAs to promote parent-child co-engagement. This approach not only boosts learning outcomes but also refines the child’s experience with the technology, addressing challenges such as imperfect speech recognition [1] and limited dialogue about the child’s personal life [39]. Additionally, considering that many bilingual children come from bi-cultural backgrounds, it is essential to situate CA conversations within culturally engaging reading materials. Children engage more effectively with content that resonates with their cultural beliefs and practices [14].

To support shared reading among bilingual children and their parents, we designed a bilingual CA and embedded it into two culturally relevant e-books. In partnership with the Joan Ganz Cooney Center at Sesame Workshop, our CA is embodied in Rosita, a well-known bilingual Sesame Street character from Mexico. Rosita not only interacts with children through a series of questions related to the stories but also suggests family questions that prompt parents to talk with their children about family experiences relating to the story. A week-long home deployment study with five Hispanic families was then carried out to evaluate the usability and effectiveness of our system.

This project represents a pilot effort to explore the long-term usability and effectiveness of a bilingual CA for facilitating parent-child shared reading in home environments. It presents a compelling case for designing a bilingual CA to actively engage families and enhance learning. Additionally, our study contributes important insights for the development of AI systems aimed at supporting collaborative reading and learning between parents and children.

2 RELATED WORK

2.1 Conversational Agents and Children’s Language Development

Children are increasingly engaging with conversational agents (CAs) through various mediums, including virtual assistants like Apple Siri and Google Assistant [28], voice-enabled tablet apps [6], Internet-connected toys [13], and social robots [33]. These CAs, serving as children’s dialogue partners during dialogic reading, have been shown to enhance children’s engagement and language learning. Smart speakers, for instance, have been used to narrate stories, ask interactive questions, and provide feedback, thus improving children’s story comprehension [38, 40]. Despite these advancements, challenges remain due to limitations in speech recognition and natural language understanding. Voice interfaces often misinterpret children’s speech [1], especially for non-native English speakers [37], and struggle with child-specific language features (e.g., pitch, verbal intonations, and typical childhood disfluencies). While there are efforts to enhance algorithm accuracy for these demographics [24], CAs still find it difficult to interpret free-form responses and engage in highly personalized conversations with children.

2.2 Enhancing Bilingual Language Development through Shared Reading

Shared reading, an integral family practice worldwide, significantly contributes to children’s language development, especially through dialogic reading where parents engage children in story-related, back-and-forth conversations [34, 36, 42]. This approach is beneficial for both monolingual and bilingual children [20, 32], utilizing varied questioning styles such as open-ended questions for elaborate discussions and close-ended ones for testing comprehension [11]. Crucially, shared reading in both languages is indispensable for children’s bilingual language proficiency, balancing the development of their dominant and less proficient languages [29]. This balanced bilingual growth is essential for children’s school readiness and maintenance of heritage language [16]. However, challenges arise when children lack sufficient exposure to both languages, often due to parents’ limited proficiency in one language [18]. This underscores the need for enriching and high-quality bilingual shared reading opportunities. In addition, the lack of culturally engaging bilingual content, which is crucial for amplifying bilingual growth [27], underlines the importance of shared reading experiences that resonate with Latine culture and families’ “funds of knowledge” [25], thereby boosting children’s learning engagement and outcomes.

3 SYSTEM DESIGN

We designed our bilingual CA, Rosita, along with the e-books in which she is embedded, with the specific goal of promoting parent-child joint engagement in reading. “Rosita Reads With My Family” includes three key design features. First, we incorporated cultural practices deemed relevant to our target participants. The stories we chose for developing the e-books are centered on Latine food culture [19], *familismo* [15], and birthday traditions [30]. Second, we aimed to support bilingual reading and conversation. The CA empowering

Rosita is capable of comprehending and carrying out dialogue in English, Spanish, and a mix of both. Families are also allowed to choose either Spanish or English as their preferred language for the printed text and narrated audio. Third, we aimed to promote parent-child joint engagement. Rosita facilitates the co-reading experience by asking two questions after each page. These question pairs are specifically designed to encourage children to reflect on and express their understanding of the story while also encouraging dialogue between the parent and the child.

3.1 Design of Rosita’s Conversations

Our CA, Rosita, employs two types of questions to facilitate children’s story comprehension and prompt parent-child discussion. The first, aimed at the child, focuses on immediate comprehension. For example, Rosita might ask, “Why does Andy want a gigantic *piñata*?” and provide tailored feedback based on the child’s response. The second, a “family question,” involves both parent and child, encouraging them to relate personal experiences to the story, such as “Please ask each other what kind of *piñata* you’d like to have at your party.” When parents and children finish the discussion, they can press anywhere on the screen to proceed to the next page. This approach prompts parents to actively engage in elaborate discussions with children during shared reading.

Rosita’s speech script was developed iteratively. Bilingual researchers in our team first developed the bilingual conversation scripts based on the stories. Next, involving 20 Latine children bilingual in English and Spanish, we conducted sessions where an experimenter read them a story and asked various questions. These sessions were recorded and analyzed to ensure the questions were clear and answerable, leading to minor adjustments. Questions that confused the children or elicited limited responses were eliminated.

3.2 Developing the Bilingual CA powering Rosita

For every question directed at the child, Rosita offers automatic feedback responsive to the child’s specific answer. With Google’s speech-to-text and DialogFlow API, the CA transcribes auditory input and classifies it by semantic intent, a process termed *natural language understanding* (NLU). As depicted in figure 2, the NLU was refined using Google’s generic pre-trained model and specialized with phrases from our e-books, ensuring precise intent detection. Considering that children can respond to any question in a variety of ways, the CA was trained to associate more than one semantic intent with each question. For instance, correct answers to “How did Sebastian feel when he saw the *piñata*?” could include “surprised”, “happy”, or “excited.”

To provide scaffolding for children’s learning, a fallback intent activates when a child does not respond or provides a response that does not match predefined intents. The CA modifies the question by using more accessible language, thus making the question less challenging. For example, the CA can reword an open-ended query to a choice-based one if the child fails to answer it (e.g. from “What is a *piñata* filled with?” to “A *piñata* is a colorful decoration. Is it filled with candy or spaghetti?”). If a fallback is still triggered, the agent provides a neutral response, revealing the correct answer and adding further explanation.

To accommodate bilingual families, we designed NLU modules for both English and Spanish intent classification to accommodate bilingual families. A multilingual detection feature first determines the language of each utterance and processes it using the corresponding NLU module. If a mixed-language response is given, the agent analyzes it with both NLU modules and replies based on the most confidently identified intent. This design caters to the linguistic flexibility demonstrated by bilingual children.

4 HOME STUDY

We carried out a week-long study in the homes of five Latine families to explore three key aspects: (i) the usage patterns and communication behaviors of both children and parents, (ii) the learning outcomes for the children, and (iii) the families’ perceptions of Rosita.

4.1 Participants

We recruited five families from the Western USA region. Each family included at least one child, aged between 4 and 6 (mean = 5.20, SD = 0.84), who was identified as the primary user of the system under study. All participants self-identified as Latine. Detailed demographic data for the participants are presented in Appendix A. Each family received \$100 as compensation.

4.2 Procedure

Our study involved two home visits, a week apart. On the first visit, families received a Surface Go tablet with our system and an introductory video. We suggested reading the e-books in both languages at least once during the week while encouraging more frequent use. Children’s knowledge of target vocabulary and parents’ demographic and home language information were collected through initial assessments and surveys. In the second visit, children’s story comprehension and vocabulary were reassessed, and parents were interviewed about their experiences, with interviews conducted in their preferred language. Further methodological details are available in the study’s Appendix B

5 RESULTS

5.1 How did families interact with Rosita?

5.1.1 Overall Usage Patterns. On average, families read 7.8 times per week (SD = 3.27, range = [5,13]), more than what we suggested. Regarding the duration of each reading session, families spent an average of 15 minutes per session (SD = 7.8, range = [1.5, 27.2]). Thus, in general, families engaged with the e-books from the beginning to the end during a usage session. Notably, there was considerable variation within each family. Families sometimes read only a few pages before concluding the session, while other times, they had more extended sessions with rich conversations and discussions.

5.1.2 Language Production. In our study, we recorded 4,039 utterances across 1,243 conversational turns from children and parents. Children responded to all of Rosita’s questions in each session, sometimes with parental assistance. Notably, children participated in more conversational turns during family questions than child questions, with an average of 1.51 turns with Rosita (SD = 0.50) compared to 4.85 turns with parents (SD = 4.92) for family questions.

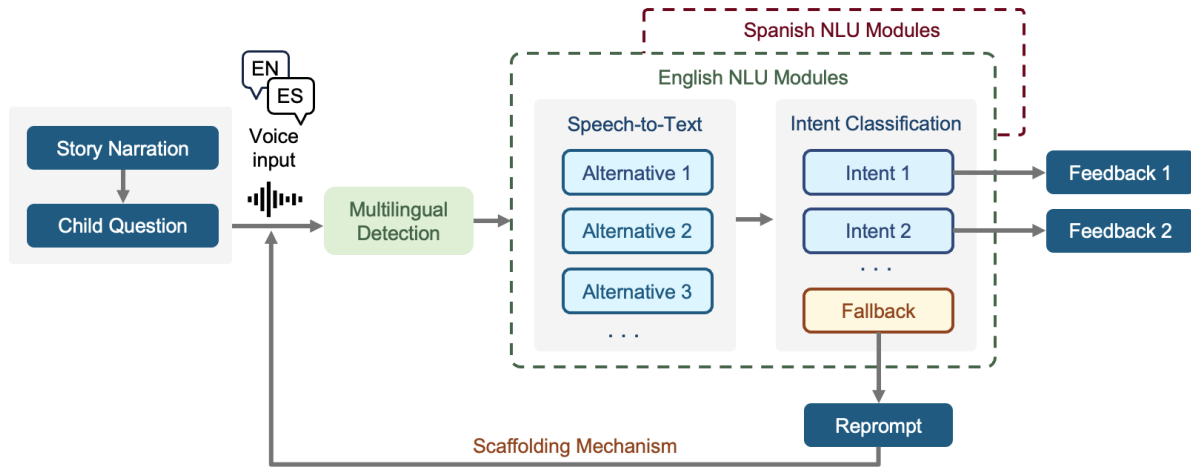


Figure 2: The bilingual natural language understanding process powering Rosita’s child questions

This indicates that family questions provided additional opportunities for the parent and the child to verbally engage in shared reading.

5.1.3 Bilingual Language Use. Families showed a preference for co-reading with Rosita in English over Spanish. Among the 39 sessions conducted with five families, 24 were in English and 15 in Spanish, mirroring their predominant home language usage as shown in table A. Notably, families who previously seldom read in one language began reading more in that language with our system. In English sessions, parents and children spoke in English 92% and 95% of the time, respectively. However, in Spanish sessions, there was a significant language blend, with parents and children using Spanish 41% and 58% of the time, respectively, and the remainder being either English or a mix of both languages within a single utterance.

5.1.4 Family Participation. Children and parents (mostly mothers, sometimes fathers) took part in the co-reading activities, but it was not uncommon for siblings to also participate in the sessions. Each member’s participation rate was determined by dividing the number of utterances they made by the total number of utterances made by all participants. We found that children were the major participants in the child-directed questions (average 73%), with parents contributing occasionally (average 23%) and siblings making up the remaining 5%. During family questions, however, the roles were more balanced, with parents participating about 52% of the time, children 39%, and siblings 9%. Siblings in families with more than one child, were more active in family questions, sometimes responding to younger siblings or extending discussions. In the absence of parents, siblings frequently interacted with each other, with the older sibling often facilitating the conversation.

5.2 What did children learn from co-reading with Rosita?

Children exhibited a solid grasp of the story’s content, as evidenced by their performance on comprehension questions in the post-test,

answering on average 86% of the questions correctly (SD = 10%, Range = [71%, 100%]). They also showed noticeable improvement in vocabulary. Before they engaged with the e-books, the children’s average accuracy rate in responding to questions about bilingual vocabulary words in the books stood at 39% (SD = 15%, Range = [16%, 52%]). Following the one-week reading experience with Rosita, their performance improved substantially, with the average accuracy rate rising to 55% (SD = 19%, Range=[27%, 80%]). Moreover, it is noteworthy that *each individual child* within the cohort exhibited discernible knowledge gains in their knowledge of vocabulary words in the e-books, with their improvements in accuracy rate ranging from 4% to 32% (Mean = 16%, SD = 12%).

5.3 How did parents and children perceive co-reading with Rosita?

We conducted a thematic analysis of parent interviews and identified five major themes regarding how parents perceived “Rosita Reads with My Family”, as detailed in this section. First, all five parents acknowledged that Rosita’s contingent interactivity helped children to engage in the reading. For example, P4 said that the child “was more paying attention because he knew that they were going to ask him a question.” The visual effects—most notably the light bulb that lit up after each correct answer—served as positive reinforcement, motivating children to read and interact with Rosita because “she (the child) felt like Rosita was talking to her. (P2)” Additionally, children’s familiarity with Rosita and Sesame Street contributed to their engagement because “it’s very approachable. It’s something that kids know. (P4)”

Second, parents noted that not only their children but also they themselves gained knowledge from reading sessions with Rosita. Parents enjoyed Rosita’s bilingual capabilities as a means of reinforcing prior knowledge and exposing children to new vocabulary across languages. For instance, P1 stated that it was helpful that Rosita “went back and said *abuela* is *grandma*.” Additionally, parents also mentioned how Rosita’s scaffolding mechanism facilitates learning. P4 mentioned that “[H]e (the child) understood a little

bit better” when Rosita provided multiple options for the child to choose from. Notably, parents reported learning shared reading strategies modeled by Rosita. P2 said, “We really like Rosita asking [family] questions at the end of each page . . . I think we’ll do more family discussion now about books, like how this is relevant to our lives and the world.”

Third, parents valued the opportunity to read and talk in both languages. P2 stated that “reading both the English and the Spanish reinforces her [child’s] understanding”. The bilingual exposure also fostered parents’ awareness of their children’s bilingual proficiency. P1 “was really shocked” when she realized “[my] kid not knowing Spanish all the way.” Parents also mentioned that the children’s limited proficiency caused challenges in their communication with Rosita. P3, for example, attributed Rosita’s failure in registering her child’s responses in Spanish due to the child’s accent.

Fourth, parents reported that Rosita fostered bonding in the family through family questions. P4 described how Rosita promoted conversation around shared memories as her child “did that with my grandma and my grandpa.” Rosita also created moments of family bonding through future-oriented discussions. P4 liked the question about what they would like to grow in the garden because “[my child] actually does spend a lot of time in the garden with my parents, so I thought that was nice because it’s kind of personal to us.” In addition to fostering these discussions, Rosita promoted parental facilitation of shared reading. P4 mentioned that Rosita’s family questions encouraged them to take the role of “an interviewer” and learn the child’s perspective about the books.

Fifth, parents appreciated that Rosita demonstrated cultural relevance through the representation of everyday Latine practices. P2 stated how her daughter enjoyed co-reading with Rosita because “she likes cooking . . . We liked how [Rosita and Elmo] wrote down the ingredients before they went grocery shopping.” Rosita also sparked conversations that connected children to previous generations. P4 stated that the cooking practices are something that her “parents taught [her] . . . something that [she] hopes to keep passing on to [her] children.” Lastly, while most parents appreciated the cultural relevance of our system, P1 asked for more diverse representations of Latine culture, “not just Mexican . . . We do want to support other Latinos”.

6 DISCUSSION

6.1 Designing Family-Oriented Conversational Agents for Shared Reading

Our project, “Rosita Reads With My Family,” extends the capabilities of CAs beyond typical one-on-one child interactions by facilitating meaningful dialogues between parents and children. For one, this CA guided children through a question-answer-feedback sequence, commonly used in educational [10] and parental settings [2], leveraging conversational technologies such as NLU for automated feedback. We found that this approach promoted children’s word learning and story comprehension by actively engaging them in the learning process. For another, the CA introduced open-ended discussion prompts to encourage family participation, aligning with the HCI community’s focus on family-centered design [4]. These prompts fostered meaningful family discussions, thereby enhancing oral language for children and strengthening parent-child bonds,

as demonstrated in our utterance analysis and parent interviews. Together, the dual functionality of our CA highlights key design considerations for future CAs –combining child-directed structured dialogues and free-flowing family discussions to balance guided learning with enriched, family-centric conversations.

Similar to many conversational agents in the field, our CA relies on real-time API requests for speech recognition and intent classification, necessitating a stable internet connection for feedback. However, as revealed in parent interviews, unstable connections leading to delayed responses can cause frustration, particularly for children. To enhance user experience, future CA designs might benefit from a proactive feedback mechanism, where the CA automatically provides neutral feedback if processing delays are detected. This would help sustain conversation flow and prevent confusion or disengagement during interaction interruptions.

6.2 Designing Conversational Agents for Bilingual Communities

We designed “Rosita Reads With My Family” to meet the bilingual and bi-cultural needs of the families, a critical factor in enhancing engagement and learning during reading. Shared reading in both English and Spanish is crucial for developing Latine bilingual children’s language skills, providing a balance between their dominant and less proficient languages [29]. A common challenge, however, is children’s inadequate exposure to both languages, often due to parental limitations in one language [3]. Our system addresses this by allowing families to choose and switch between English and Spanish reading, featuring a bilingual CA that understands and responds in both languages. This flexibility proved beneficial in our study, as families who previously seldom read in one language began reading more in that language with our system. Parents also showed positive perceptions about the bilingual learning opportunities. Moreover, the bilingual functionality of the CA matched the linguistic flexibility of our participants, who often used a mix of both languages in their communication.

Regarding the cultural aspect, it is well-acknowledged that children learn more effectively when the learning experiences align with their cultural backgrounds [17]. In our project, we incorporated Rosita, the bilingual CA, into e-books that incorporated Latine cultural practices and values. Parents appreciated this cultural relevance, noting that it facilitated meaningful conversations and helped familiarize their children with Latine cultural themes. This approach also encouraged children to actively use and engage with the Spanish language, a positive outcome highlighted by the parents. Therefore, future research on CAs should focus not only on their bilingual linguistic capabilities but also on situating the CA’s conversations within learning materials relevant to the cultural practices and beliefs of bilingual families, thereby enhancing motivation and engagement.

However, parents noted that the Spanish version of our CA was less effective in registering children’s responses compared to the English version, attributed to the children’s limited Spanish skills and distinct accents. This highlights a broader issue with NLU systems, which are often trained on adult speech and primarily in English [23], leading to less effective processing of children’s

speech and other languages [1, 37]. To enhance these systems, future research should focus on diversifying training datasets with speech samples from children and speakers of various languages and dialects. Furthermore, parents' feedback regarding the representation of Latine culture in our e-books suggests a need for more diverse and authentic cultural content. Current materials on Latine cultures focused predominantly on Mexican culture do not fully represent the broader Latine experience. Addressing this, future efforts should involve co-design strategies with families, ensuring learning materials accurately and inclusively reflect the diverse perspectives and experiences of the Latine community.

6.3 Future Direction

There are two directions for future work to improve the design and provide empirical evidence regarding how AI can support shared reading. One direction for future research is to explore the comparative benefits of using CAs for solo child engagement versus parent-child co-reading. Numerous studies indicate that involving parents in reading leads to children's exposure to richer language and higher engagement levels (e.g., [38]). It is plausible that parent-child co-reading could offer additional benefits over children reading independently. To investigate this, future studies could employ a randomized controlled trial with a larger sample size to contrast these two approaches, assessing their impact on children's language, exposure, learning, and engagement.

Another avenue for future research direction could involve integrating adaptive learning algorithms CAs for a personalized reading experience tailored to each child's language skills and learning pace. The language proficiency of bilingual children can significantly differ both between languages and among individuals [9]. This approach would allow CAs to adjust language complexity, questions, and feedback in real-time based on the child's responses, offering more effective support for children's bilingual language development.

7 CONCLUSION

We conducted a one-week home study of a bilingual CA embedded in e-books to facilitate parent-child shared reading, with the aim of enriching parent-child interactions and supporting Latine children's bilingual language growth. Our study suggested that the bilingual CA successfully engaged children and parents in shared reading for an extended period of time, elicited rich verbal interactions between them, and supported children's language learning. Our study demonstrated that conversational technologies tailored to families' social and cultural contexts are promising in enriching families' home language environment in the long run.

ACKNOWLEDGMENTS

This paper is based upon work supported by the National Science Foundation under Grant No. 2115382. We thank the Joan Ganz Cooney Center at Sesame Workshop for collaboration on this research. We thank the Child Studies Collaborative for assisting with participant recruitment. We thank the families who participated in this study. We also thank the undergraduate research assistants who assisted with data coding.

REFERENCES

- [1] Erin Beneteau, Olivia K Richards, Mingrui Zhang, Julie A Kientz, Jason Yip, and Alexis Hiniker. 2019. Communication breakdowns between families and Alexa. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.
- [2] Pamela Blewitt, Keiran M Rump, Stephanie E Shealy, and Samantha A Cook. 2009. Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology* 101, 2 (2009), 294.
- [3] Krista Byers-Heinlein. 2013. Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism: Language and Cognition* 16, 1 (2013), 32–48.
- [4] Bengisu Cagiltay, Rabia Ibtasar, Joseph E Michaelis, Sarah Sebo, and Bilge Mutlu. 2023. From Child-Centered to Family-Centered Interaction Design. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference*. 789–791.
- [5] Justine Cassell. 2001. Embodied conversational agents: representation and intelligence in user interfaces. *AI magazine* 22, 4 (2001), 67–67.
- [6] Yi Cheng, Kate Yen, Yeqi Chen, Sijin Chen, and Alexis Hiniker. 2018. Why doesn't it work? Voice-driven interfaces and young children's communication repair strategies. In *Proceedings of the 17th ACM conference on interaction design and children*. 337–348.
- [7] Victoria Clarke, Virginia Braun, and Nikki Hayfield. 2015. Thematic analysis. *Qualitative psychology: A practical guide to research methods* 3 (2015), 222–248.
- [8] Graham Crookes. 1990. The utterance, and other basic units for second language discourse analysis. *Applied linguistics* 11, 2 (1990), 183–199.
- [9] Caroline De Lamo White and Lixian Jin. 2011. Evaluation of speech and language assessment approaches with bilingual children. *International Journal of Language & Communication Disorders* 46, 6 (2011), 613–627.
- [10] Richa S Deshmukh, Jill M Pentimonti, Tricia A Zucker, and Bridget Curry. 2022. Teachers' use of scaffolds within conversations during shared book reading. *Language, Speech, and Hearing Services in Schools* 53, 1 (2022), 150–166.
- [11] Richa S Deshmukh, Tricia A Zucker, Sherine R Tambyraja, Jill M Pentimonti, Ryan P Bowles, and Laura M Justice. 2019. Teachers' use of questions during shared book reading: Relations to child responses. *Early Childhood Research Quarterly* 49 (2019), 59–68.
- [12] Rebecca A Dore, Brenna Hassinger-Das, Natalie Brezack, Tara L Valladares, Alexis Paller, Lien Vu, Roberta Michnick Golinkoff, and Kathy Hirsh-Pasek. 2018. The parent advantage in fostering children's e-book comprehension. *Early Childhood Research Quarterly* 44 (2018), 24–33.
- [13] Stefania Druga, Randi Williams, Hae Won Park, and Cynthia Breazeal. 2018. How smart are the smart toys? Children and parents' agent interaction and intelligence attribution. In *Proceedings of the 17th ACM Conference on Interaction Design and Children*. 231–240.
- [14] Geneva Gay. 2018. *Culturally responsive teaching: Theory, research, and practice*. teachers college press.
- [15] Sandra M Gonzales. 2019. Cultivating familismo: Belonging and inclusion in one Latina/o learning community. *International Journal of Inclusive Education* 23, 9 (2019), 937–949.
- [16] Martin Guhn, Constance Milbrath, and Clyde Hertzman. 2016. Associations between child home language, gender, bilingualism and school readiness: A population-based study. *Early Childhood Research Quarterly* 35 (2016), 95–110.
- [17] Carol Scheffner Hammer and Brook Sawyer. 2016. Effects of a culturally responsive interactive book-reading intervention on the language abilities of preschool dual language learners: A pilot study. *HS Dialog: The Research to Practice Journal for the Early Childhood Field* 19, 2 (2016).
- [18] Erika Hoff and Cynthia Core. 2013. Input and language development in bilingually developing children. In *Seminars in speech and language*, Vol. 34. Thieme Medical Publishers, 215–226.
- [19] Zilkia Janer. 2008. *Latino food culture*. ABC-CLIO.
- [20] Catriona Kennedy and Aoife McLoughlin. 2022. Developing the Emergent Literacy Skills of English Language Learners Through Dialogic Reading: A Systematic Review. *Early Childhood Education Journal* (2022), 1–16.
- [21] Anthony J Liddicoat. 2021. *An introduction to conversation analysis*. Bloomsbury Publishing.
- [22] Silvia B Lovato, Anne Marie Piper, and Ellen A Wartella. 2019. Hey Google, do unicorns exist? Conversational agents as a path to answers to children's questions. In *Proceedings of the 18th ACM international conference on interaction design and children*. 301–313.
- [23] Nina Markl. 2022. Language variation and algorithmic bias: understanding algorithmic bias in British English automatic speech recognition. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 521–534.
- [24] Marco Matassoni, Roberto Gretter, Daniele Falavigna, and Diego Giuliani. 2018. Non-native children speech recognition through transfer learning. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 6229–6233.
- [25] Luis C Moll, Cathy Amanti, Deborah Neff, and Norma Gonzalez. 1992. Funds of knowledge for teaching: Using a qualitative approach to connect homes and

- classrooms. *Theory into practice* 31, 2 (1992), 132–141.
- [26] Koichi Mori, Rafael Ballagas, Glenda Revelle, Hayes Raffle, Hiroshi Horii, and Mirjana Spasojevic. 2011. Interactive rich reading: enhanced book reading experience with a conversational agent. In *Proceedings of the 19th ACM international conference on Multimedia*. 825–826.
 - [27] Rahat Naqvi, Anne McKeough, Keoma Thorne, and Christina Pfitscher. 2013. Dual-language books as an emergent-literacy resource: Culturally and linguistically responsive teaching and learning. *Journal of Early Childhood Literacy* 13, 4 (2013), 501–528.
 - [28] Cansu Oranç and Azzurra Ruggeri. 2021. “Alexa, let me ask you something different” Children’s adaptive information search with voice assistants. *Human Behavior and Emerging Technologies* 3, 4 (2021), 595–605.
 - [29] Alison H Paris and Scott G Paris. 2003. Assessing narrative comprehension in young children. *Reading Research Quarterly* 38, 1 (2003), 36–76.
 - [30] Mari Riojas-Cortez and Raquel Cataldo. 2015. Using children’s literature to understand values, traditions, and beliefs within Latino Family Systems. *Multicultural literature for Latino bilingual children: Their words, their worlds* (2015), 83–100.
 - [31] Kimiko Ryokai, Cati Vaucelle, and Justine Cassell. 2003. Virtual peers as partners in storytelling and literacy learning. *Journal of computer assisted learning* 19, 2 (2003), 195–208.
 - [32] Monique Sénéchal and Laura Young. 2008. The effect of family literacy interventions on children’s acquisition of reading from kindergarten to grade 3: A meta-analytic review. *Review of educational research* 78, 4 (2008), 880–907.
 - [33] Suleman Shahid, Emiel Krahmer, and Marc Swerts. 2014. Child–robot interaction across cultures: How does playing a game with a social robot compare to playing a game alone or with a friend? *Computers in Human Behavior* 40 (2014), 86–100.
 - [34] Timothy Shanahan and Christopher J Lonigan. 2010. The National Early Literacy Panel: A summary of the process and the report. *Educational Researcher* 39, 4 (2010), 279–285.
 - [35] Guadalupe Valdés. 2005. Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized? *The modern language journal* 89, 3 (2005), 410–426.
 - [36] Grover J Whitehurst, Francine L Falco, Christopher J Lonigan, Janet E Fischel, Barbara D DeBaryshe, Marta C Valdez-Menchaca, and Marie Caulfield. 1988. Accelerating language development through picture book reading. *Developmental psychology* 24, 4 (1988), 552.
 - [37] Yunhan Wu, Daniel Rough, Anna Bleakley, Justin Edwards, Orla Cooney, Philip R Doyle, Leigh Clark, and Benjamin R Cowan. 2020. See what I’m saying? Comparing intelligent personal assistant use for native and non-native language speakers. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*. 1–9.
 - [38] Ying Xu, Joseph Aubele, Valery Vigil, Andres S Bustamante, Young-Suk Kim, and Mark Warschauer. 2022. Dialogue with a conversational agent promotes children’s story comprehension via enhancing engagement. *Child Development* 93, 2 (2022), e149–e167.
 - [39] Ying Xu, Stacy Branham, Xinwei Deng, Penelope Collins, and Mark Warschauer. 2021. Are Current Voice Interfaces Designed to Support Children’s Language Development?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–12.
 - [40] Ying Xu, Kunlei He, Valery Vigil, Santiago Ojeda-Ramirez, Xuechen Liu, Julian Levine, Kelsyann Cervera, and Mark Warschauer. 2023. “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*. 160–172.
 - [41] Ying Xu, Dakuo Wang, Penelope Collins, Hyelim Lee, and Mark Warschauer. 2021. Same benefits, different communication patterns: Comparing Children’s reading with a conversational agent vs. a human partner. *Computers & Education* 161 (2021), 104059.
 - [42] Andrea A Zevenbergen and Grover J Whitehurst. 2003. Dialogic reading: A shared picture book reading intervention for preschoolers. *On reading books to children: Parents and teachers* 177 (2003), 200.

A TABLE OF PARTICIPANT DEMOGRAPHICS

B RESEARCH METHODS

B.1 Families’ Usage Patterns and Communication Behavior

To obtain the **usage patterns**, we calculated how many times each family read with Rosita and the duration of each usage session, based on the log data saved in the devices. For a usage session to be included in the analytical sample, a family had to read at least one page of the e-book in that session. To examine the quantity and quality of families’ **language production** during the reading

sessions, we coded the transcriptions of all the reading sessions. Note that Rosita asks one question for the child and suggests a discussion prompt for the family after each page. Each of these questions/prompts marked the beginning of a coding unit, and each coding unit included all parent or child utterances made in response to the question but prior to Rosita’s subsequent question. An utterance is a continuous unit of speech, beginning and ending with a pause [8]. We used the number of conversational turns and mean length of utterances (MLU) to capture the quantity of language production. A conversational turn is defined as a back-and-forth verbal exchange between the interlocutors [21], while MLU is the number of words in each utterance. We also coded which language the speaker used for each of their utterances to capture their **bilingual language use**. Additionally, we coded which family members participated in the reading and to what extent they were involved to understand **family participation**.

B.2 Children’s Learning Outcomes

We asked children questions about the story content and target vocabulary words in the two e-books. Regarding **story comprehension**, we asked children 7–8 questions about each e-book. We scored children’s responses as two points if they answered the question correctly. If they failed to answer the question, we provided them with three choices. They got one point if they answered it correctly in the second trial.

The **vocabulary items** consisted of two types: vocabulary recognition (receptive) and vocabulary definition (expressive). For vocabulary recognition, children were tasked with selecting the appropriate picture corresponding to a given word. Responses were scored as one point if they successfully recognized the vocabulary, and zero if not. For vocabulary definition, responses were scored from zero to two, based on completeness. Each response was scored independently by the two principal authors according to a pre-established scoring guide. The raters then met to discuss and resolve discrepancies.

B.3 Parents’ and Children’s Perceptions

We conducted a thematic analysis of parent interviews [7] to identify how they perceive “Rosita Reads with My Family.” Two authors first generated succinct codes that captured the major aspects of the interview questions. They then collated interview quotes relevant to each code. An inter-rater reliability of 0.90 was reached among them. Lastly, they examined the quotes under each code to further develop themes by splitting, combining, or discarding them, which formed the final themes and sub-themes.

Family	Home Language		Home Literacy Environment				Digital Reading	CA ³ Exposure
	Parent to Child	Child to Parent	Spanish		English			
			Reading ¹	Books ²	Reading	Books		
F1	Equally Eng & Spa	Mostly Eng	3 days 10-20 min	6-10 books	7 days 10-20 min	20+ books	Weekly	Weekly
F2	Mostly Spa	Mostly Eng	3 days 0-10 min	20+ books	6 days 10-20 min	20+ books	Weekly	Weekly
F3	Mostly Eng	Only Eng	1 days 0-10 min	1-5 books	5 days 20-30 min	20+ books	Never	Daily
F4	Equally Eng & Spa	Mostly Eng	1 days 0-10 min	1-5 books	5 days 10-20 min	20+ books	Never	Monthly
F5	Mostly Eng	Mostly Eng	1 days 0-10 min	1-5 books	1 days 10-20 min	1-5 books	Monthly	Never

Table 1: Demographics of participants

Notes:

¹ The number of days per week families read in the specified language and the average duration for each time they read.

² The number of children's books in the specified language in the household.

³ CA refers to the CA.