"I Said Knight, Not Night!": Children's Communication Breakdowns and Repairs with AI Versus Human Partners

Zhixin Li
University of Michigan
Ann Arbor, USA
zhixinli@umich.edu

Trisha Thomas
University of Michigan
Ann Arbor, Michigan, USA
ttrisha@umich.edu

Chi-Lin Yu
University of Michigan
Ann Arbor, Michigan, USA
chilinyu@umich.edu

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

ABSTRACT

In this study, we explored communication breakdowns and repair strategies among 71 children aged 4-8 years while co-creating stories with a generative AI agent enabled by Large Language Models and a human partner. Analyzing approximately 1420 minutes of video recordings, our findings reveal that children experienced more communication breakdowns when interacting with the AI partner but attempted repairs more frequently with human counterparts. Notably, children who attributed greater mind perception to non-human entities were more proactive in attempting repairs during interactions with both human and AI partners, with this trend being more pronounced when children interacted with AI. This work-in-progress offers theoretical contributions by illuminating the interplay between perception and communication. It also underscores important design considerations for developing LLM-enabled generative AI agents that are socio-cognitively responsible and aligned with children's perceptions.

CCS CONCEPTS

• **Human-centered computing → Empirical Studies in HCI, Social and professional topics → Children**.

KEYWORDS

Conversational AI; Generative AI; Large Language Models; Communication breakdowns; Repair strategies; Mind perception

ACM Reference Format:

Zhixin Li, Trisha Thomas, Chi-Lin Yu, and Ying Xu. 2024. "I Said Knight, Not Night!": Children's Communication Breakdowns and Repairs with AI Versus Human Partners. In *Interaction Design and Children (IDC '24), June 17–20, 2024, Delft, Netherlands.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3628516.3659394

IDC '24, June 17–20, 2024, Delft, Netherlands

ACM ISBN 979-8-4007-0442-0/24/06
https://doi.org/10.1145/3628516.3659394

1 INTRODUCTION AND BACKGROUND

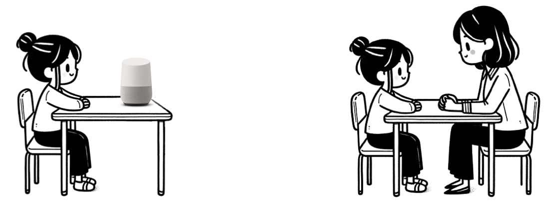
Engaging in conversations with others is crucial for young children as it enables their social interactions and allows them to acquire information that develops their understanding of the world. Yet, one important skill children need to learn during these conversations is to navigate through communication breakdowns, in which children and their conversation partners misunderstand each other or fail to continue the flow of conversations. Research has consistently shown that even young children are able to use a range of repair strategies [2, 6, 8, 12, 20], such as repetition, modification, or pivoting the direction of the conversation. Children become more sophisticated in implementing these strategies as they grow older [6, 20], likely due to their developing language proficiency and socio-cognitive skills such as perspective-taking [26].

As children become increasingly likely to engage in conversations not only with humans but also with artificial intelligence (AI) such as Siri, Alexa, and ChatGPT, the issue of communication breakdown may be further amplified. It is not uncommon for these AI agents to mistranscribe children's speech from one word to another or misinterpret their intentions [23]. Indeed, previous studies have uncovered that children did encounter breakdowns during their interactions with AI agents and that children demonstrated persistence and implemented similar repair strategies as in human-to-human interactions [4, 8, 10, 22–24, 36]. There was also evidence that children may require additional support from adults to handle such breakdowns [4, 8, 34, 40].

However, the extent to which children's motivation to repair communication breakdowns might differ when interacting with AI compared to humans remains unclear. Theoretically, children's repair behaviors could be influenced by their perceptions of AI, such as whether they perceive AI as capable of experiencing emotions or sensations [3, 11, 13, 37], or possessing agency to dictate its own behaviors [31]. Studies concerning preschool-aged children interacting with social robots [5, 21, 27, 32] have suggested that such mind perception tendency [16] is linked to more social-oriented and more effortful interactions with technologies.

This current study aims to bring these strands together and provide an in-depth examination of children's breakdown and repair strategies with both AI and human partners, as well as explore the role that children's general mind-perception tendency (i.e., the tendency to perceive whether an entity has a mind or not) might play in their communicative behaviors. To this end, our study involved 71 children aged four to eight years, who were randomly assigned to co-create stories with either an AI agent, developed using OpenAI's GPT-4, or a human partner. Drawing on this data, our study addresses three questions:

- **RQ1:** Did children encounter more breakdowns while conversing with AI versus humans? What types of breakdowns did they encounter?
- **RQ2:** Were children less likely to repair breakdown while conversing with AI versus humans? What types of repair strategies did children use?
- **RQ3:** How does children's general mind-perception tendency affect their repair behaviors in communication breakdowns with an AI versus a human partner?



(a) AI condition **(b) Human condition**

Figure 1: Story co-creation session setup

2 THE DESIGN OF STORY CO-CREATING SESSION

The conversation in which children engaged, either with an AI or a human partner, occurred in the context of storytelling. We selected storytelling because it is a creative task that allows the conversation to be child-driven and grounded in the child's lived experience, which might simulate social-oriented responses from children while showcasing the unique capabilities of generative AI as well as human creativity. Children engaged in creating two stories with their respective partner (human or AI), following a similar structure. For the AI condition, we utilized OpenAI's GPT-4, the most advanced large language model available at the time. For details on the design of the AI storyteller, see [39].

The storytelling session involved turn-based dialogue between the participant and their partner. See Figure 1 for the setup of our story co-creation session. The story creation process was structured such that the partner first asked the children to decide on the protagonist and setting for their story. The partner then incorporated this information into the narrative. Once the setup was provided, the partner began with the introductory paragraph, followed by questions encouraging children to consider and articulate character emotions and plot developments. The partner acknowledged the children's input, continued the story with a brief 1-2 sentence addition, and then encouraged the children to expand on their previous contribution to further the narrative. In the human condition, an experimenter sat next to the child while engaging in conversation. In the AI condition, the child interacted with a smart speaker device. Children were encouraged to pick their favorite plush toy and act out the stories to support their embodied interactions. The story co-creation lasted about 20 minutes.

3 METHOD

3.1 Participants

A total of 71 children, aged 4-8 years (mean age = 6.35, 62.50% female), participated in our study. 61.43% of children were self-identified as White, 18.57% of children were mixed-race, and 8.57% of them were Asian. All participants reported that they primarily spoke English at home. The majority of them reported having frequent interactions with commercial conversational agents (e.g., Siri, Alexa), with 22.86% reported daily usage (See the participant information in Appendix A.). These children were randomly assigned to two experimental conditions (AI and Human) to engage in story co-creation sessions. We carried out a balance check and confirmed that the randomization was successful; there were no significant differences between the two conditions in terms of participants' age, gender, race/ethnicity, or previous use of conversational agents.

3.2 Measures

3.2.1 General Mind-perception Tendency and Oral Language Proficiency.

To measure children's baseline level of mind-perception tendency, we administered the Individual Differences in Anthropomorphism Questionnaire - Child Form [29] wherein children were asked to indicate their attribution of mind to three categories of entities: animals, technologies, and nature. Example questions include "Does a robot know what it is?", "Does a cheetah have feelings?", and "Does a tree think for itself?" If a child provided an affirmative response, they were then asked to indicate the extent to which they agreed on a three-point scale. This instrument showed satisfactory validity and reliability among children aged three to six to reflect children's general tendency to perceive whether an entity has a mind.

To measure children's baseline oral language proficiency, we used a widely used paradigm where children were asked to generate narratives based on a wordless picture book [14]. Video recordings of children's generated stories were transcribed verbatim and then scored by summing up the number of narrative elements (e.g., setting, initiating problems, goal-oriented actions) contained, using a validated story grammar approach [15, 18, 30].

3.2.2 Coding of Communication Breakdown and Repair Behaviors.

We transcribed children's interactions from video recordings which formed the basis of this analysis. The unit of the coding is each "conversation turn" between a child and their partner, which consists of a conversational partner's question, the child's response, and any subsequent feedback from the partner. There were a total of 1364 conversation turns in our sample, with an average of approximately 20 turns between each child and their respective partner.

Breakdown was defined by any instance where the flow of conversation was interrupted [1]. For each conversation turn, we coded whether a breakdown occurred (1) or not (0). To further categorize the types of breakdown, we developed our protocol based on a coding system for describing conversational breakdowns in preschool children [38]. *Auditory aspect* was coded if a breakdown was due to the child's unclear utterances (e.g., 'shell' to 'sell'). *Verbal aspect* was coded when the sentence structure of utterances was overly complex or contained grammatical errors. *Contextual aspect* was coded if the child failed to provide enough context for the partner to understand the message (e.g., "Can I take 'that' home?"). Additionally, we added an *engagement aspect* to capture breakdowns that occurred due to the child's distraction or lack of engagement, such as when the child was distracted by surrounding noise.

Repairs were defined as any attempt by the child to address a breakdown. For each turn, we coded whether a repair was initiated by the child (yes as 1; no as 0). To further capture children's types of repair strategies, we developed our coding protocol based on [6, 25]. The resulting protocol included four strategies. *Repetition* was coded when the child repeated their original utterance. *Correction* was coded if the child corrected their partner by stating what they actually meant (e.g., "I didn't say *Uni*, I said *I want a Pokémon*. Did you hear me?"). *Request for clarification* was coded if the child explicitly expressed their confusion resulting from a breakdown and/or asked the partner to provide further information. *Accommodation* was coded if the child adjusted responses to accommodate misinterpretations (e.g., AI: "Oh, a hat *[The child had said "Perhaps"]*. That sounds fun. Can you tell me what kind of hat it is? What color is it?" Child: *[Awkwardly smiled]* "Yellow."). Two coders separately coded the same transcriptions and the inter-rater reliability (IRR) for coding consistency was computed, yielding an IRR of 88.5%. See complete coding protocol in Appendix B.

4 PRELIMINARY FINDINGS

4.1 Communication Breakdowns with AI vs Humans

The frequency of children's breakdown by conditions was presented in Figure 2. Descriptively, children interacting with the AI experienced more breakdowns than those interacting with humans. To statistically examine whether breakdowns occur more frequently when children interact with AI compared to humans (RQ1), we conducted a generalized linear mixed-effects model (GLMM), controlling for children's age and baseline oral language proficiency.

This analysis confirmed children interacting with AI were more likely to experience breakdowns. Condition was a significant predictor of communication breakdowns, indicating a reduced likelihood during human interactions ($\beta = -0.90$, $SE = 0.30$, $z = -3.30$, $p < .001$).

Age also significantly influenced the probability of experiencing a breakdown, with older children being less likely to experience a breakdown ($\beta = -0.40$, $SE = 0.13$, $z = -3.08$, $p = .002$). Contrary to our expectations, baseline language abilities were not a significant predictor of breakdown occurrence ($\beta = 0.005$, $SE = 0.029$, $z = 0.16$, $p = .874$). The random effects analysis revealed considerable variability among children in their propensity for breakdowns (Variance = 0.64, $SD = 0.80$), suggesting significant differences in children's baseline likelihood of experiencing a breakdown.

When we further examine the different types of breakdowns children experience (Table 1), it is noticeable that children faced substantial challenges in pronunciation and speech, with auditory aspects accounting for the majority of unsuccessful communications. This was a more salient issue while interacting with AI. Verbal and engagement aspects also contributed considerably to breakdowns in both human and AI conditions.

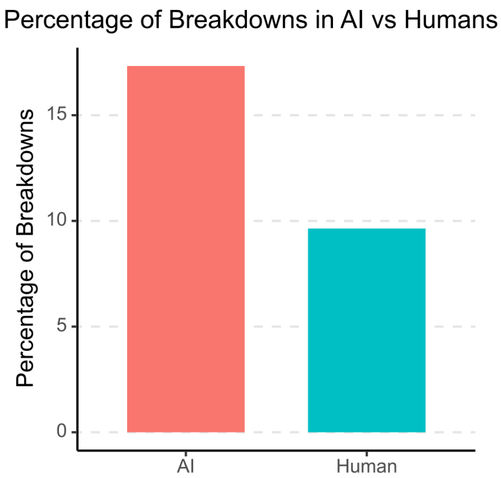


Figure 2: The bar chart displays the percentage of communication breakdowns during conversations with human and AI partners

Table 1: Distribution of communication breakdown types by study conditions

	Condition		
Breakdown Type	**AI**	**Human**	**Full**
Auditory aspect	65.71%	43.24%	56.42%
Verbal aspect	18.10%	16.22%	17.32%
Contextual aspect	2.86%	16.22%	8.38%
Engagement aspect	13.33%	24.32%	17.88%
Total	105	74	179

4.2 Repairs with AI vs Humans

Within the conversational turns where children encountered breakdowns, the frequency of their repair attempts is displayed in Figure 3. It appeared that children in the human condition were more likely to initiate a repair than those conversing with an AI. To statistically examine whether children are more likely to attempt repairs during interactions with AI compared to humans (RQ2), we employed a similar logistic GLMM. This analysis was restricted to instances of communication breakdowns among children who had experienced at least one such event ($n = 44$).

Results from the model were consistent with the descriptive figure. Interaction condition significantly predicted repair attempts ($\beta = 3.02$, $SE = 0.86$, $z = 3.50$, $p < .001$), with children more frequently initiating repair in interactions with humans than with AI. Age ($\beta = 0.09$, $SE = 0.29$, $z = 0.31$, $p = .757$) and baseline language abilities ($\beta = 0.40$, $SE = 0.06$, $z = 0.64$, $p = .526$) did not significantly predict repair behaviors. Significant variability among children (*SD* = 1.33) was noted, highlighting individual differences in the propensity for repair.

When further exploring the types of repair strategies (Table 2), it appeared that children employed different strategies depending on their conversational partner. Repetition was most frequent when engaging with human partners, while requesting for clarification was the most common strategy with AI. Interestingly, children were found to accommodate, or go along with the conversation flow, to compensate for the AI's errors, a pattern not observed with human partners.

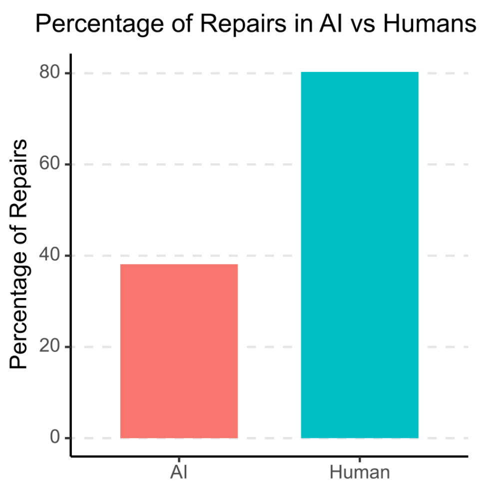


Figure 3: The bar chart displays the percentages of repair attempts in response to communication breakdowns during conversations with human and AI partners

Table 2: Distribution of repair strategies types by study conditions

	Condition		
Repair Type	**AI**	**Human**	**Full**
Repetition	8.89%	51.72%	33.01%
Correction	17.78%	24.14%	21.36%
Request for clarification	53.33%	24.14%	36.89%
Accommodation	20.00%	0.00%	8.74%
Total	45	58	103

4.3 The Moderation Role of Mind Perception Tendency in Repair Behaviors

Given our additional interest in examining how children's general mind-perception tendency might influence their repair behaviors (RQ3), we expanded our analysis to include the mind-perception scores as a moderator in a subsequent model focusing on the repair data. This approach aimed to elucidate the effect of mind-perception tendency on repair attempts and whether such effects differed between the human and AI conditions. Results showed that, overall, mind-perception tendency influenced the likelihood of repair ($\beta = 0.10$, $SE = 0.05$, $z = 2.19$, $p = .028$; see Figure 4.). This finding suggested that children's tendency to perceive whether an entity has a mind or not — regardless of whether those entities are animal, natural, or artificial entities — played a crucial role in determining whether they engaged in repair behaviors. However, the interaction between condition (Human or AI) and general mind-perception tendency was not significant ($\beta = 0.10$, $SE = 0.12$, $z = 0.86$, $p = .391$). This finding indicated that the effect of mind-perception tendency on repair likelihood did not differ substantially between different interaction partners (human or AI). In other words, children's inclination to initiate repair attempts was influenced by their general mind-perception tendency, and this influence remained consistent across interactions with both AI and humans.

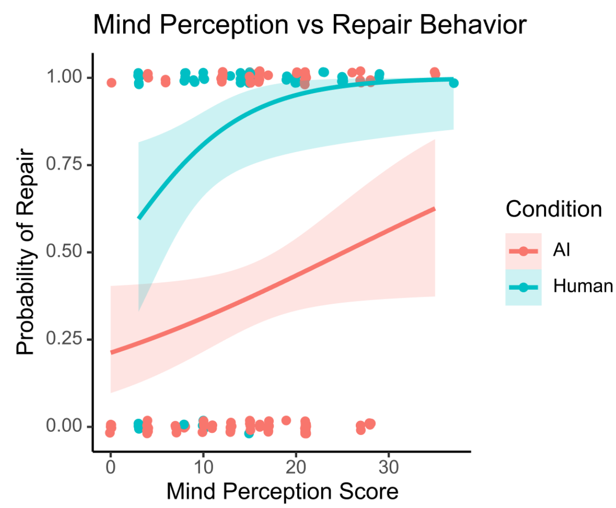


Figure 4: The logistic curve shows repair probability based on mind perception tendency

5 DISCUSSION AND FUTURE WORK

In our study, we examined how children aged 4 to 8 years addressed communication breakdowns with an AI versus humans. This comparison sheds light on the degree to which children apply human social heuristics when engaging with AI and can potentially inform the design of child-friendly AI systems.

Our first research question highlighted the inherent challenges in interactions between children and AI [4, 8, 10, 23]. It also revealed that communication breakdowns are not unique to these interactions but rather represent a universal aspect of communicative development [2, 9, 19, 28], transcending the nature of the interlocutor. Our second research question focused on children's attempts to repair those breakdowns if they occurred. The different likelihoods, as well as repair strategies utilized, are likely driven by children's perceptions or expectations of the AI's capabilities. In particular, children are much more likely to let go of AI's misunderstandings rather than persist in clarification efforts, which was not observed while children interacted with human partners. This leads to our third research question, further investigating the underlying motivations behind children's repair behaviors, particularly focusing on their tendency to perceive AI as a mental entity like humans. Interestingly, children who were more likely to attribute human-like qualities to non-human entities were generally more inclined to initiate repair when interacting with AI. Their likelihood of initiating a repair that resembles what is observed in human interactions suggests a cognitive bridging where mind perceptions influence children's communicative behaviors [33].

The emergence of accommodation as a repair strategy, uniquely observed in the context of AI interaction, merits further exploration. Follow-up analyses using additional post-test data can help disentangle the motivations behind the employment of this strategy in child-AI communication and further understand how children conceptualize the minds of their conversation partners [35]. Additionally, examining the cognitive load differences in these interactions may reveal more about our observed effects. Notably, prior research indicates a tendency to accommodate speech errors more readily from individuals perceived as belonging to an out-group, even without conscious effort [7]. Could a similar mechanism be influencing children's interactions with AI? Furthermore, the role of familiarity of AI technologies in shaping children's communication choices with AI warrants further examination. Many directions remain to be explored in understanding the nuanced interplay of communicative breakdowns and repairs in these interactions.

Moving forward, it is essential to consider the broader implications for AI design and child development. Key questions include: How might ongoing interactions with AI shape children's communication expectations and their tolerance for errors, whether from AI or humans? Furthermore, how can the insights gained from these observed behaviors guide the creation of AI systems that are more attuned to the needs of young users? As children's daily experiences increasingly include interactions with AI, the need for researchers to address these inquiries similarly grows [17]. This approach not only promises to enhance our understanding of child-AI interaction but also to inform the development of technologies that foster beneficial growth and learning. Overall, these preliminary findings underscore the importance of considering children's perceptual frameworks in the design of AI, advocating for systems that can adaptively respond to the communicative expectations children bring to these interactions.

6 PARTICIPATION AND SELECTION OF CHILDREN

Our study was approved by the Institutional Review Board of the University of Michigan. Children were recruited from two public libraries in a mid-west city in the U.S. between July and January 2024. A total of 71 Children participated in our study which included completing two sessions during story co-creation activities with either an AI partner or a human counterpart. All children consented to participate in the study and were informed that they could discontinue at any time. Upon completion of the study, children who completed the procedure were given a story book and their parents received a $20 Visa gift card for their participation.

ACKNOWLEDGMENTS

This research is supported by the National Science Foundation under Grant No. 2302730 and by the SUCCEED Faculty Grants Opportunity facilitated by the University of Michigan ADVANCE program. We thank the Ann Arbor District Library for providing space for us to carry out study sessions. We also thank the families and children who participated in our study.

REFERENCES

[1] Gabriella Airenti, Bruno G. Bara, and Marco Colombetti. 1993. Failures, exploitations and deceits in communication. *Journal of Pragmatics* 20, 4 (1993), 303–326. https://doi.org/10.1016/0378-2166(93)90055-T

[2] Dianne Alexander, Amy Wetherby, and Barry Prizant. 1997. The Emergence of Repair Strategies in Infants and Toddlers. *Semin Speech Lang* 18, 3 (1997), 197–212. https://doi.org/10.1055/s-2008-1064073

[3] Valentina Andries and Judy Robertson. 2023. Alexa doesn't have that many feelings: Children's understanding of AI through interactions with smart speakers in their homes. *Computers and Education: Artificial Intelligence* 5 (2023), 100176. https://doi.org/10.1016/j.caeai.2023.100176

[4] 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* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300473

[5] Kimberly A Brink and Henry M. Wellman. 2020. Robot teachers for children? Young children trust robots depending on their perceived accuracy and agency. *Developmental psychology* 56 (05 2020), 1268–1277. https://doi.org/10.1037/dev0000884

[6] Bonnie Brinton, Martin Fujiki, Diane F. Loeb, and Erika Winkler. 1986. Development of conversational repair strategies in response to requests for clarification. *Journal of Speech and Hearing Research* 29, 1 (1986), 75–81. https://doi.org/10.1044/jshr.2901.75

[7] Sendy Caffarra and Clara D. Martin. 2019. Not all errors are the same: ERP sensitivity to error typicality in foreign accented speech perception. *Cortex* 116 (2019), 308–320. https://doi.org/10.1016/j.cortex.2018.03.007 Structure in words: the present and future of morphological processing in a multidisciplinary perspective.

[8] 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* (Trondheim, Norway) *(IDC '18)*. Association for Computing Machinery, New York, NY, USA, 337–348. https://doi.org/10.1145/3202185.3202749

[9] Liane Comeau, Fred Genesee, and Morton Mendelson. 2007. Bilingual children's repairs of breakdowns in communication. *Journal of Child Language* 34, 1 (2007), 159–174. https://doi.org/10.1017/S0305000906007690

[10] Yao Du, Kerri Zhang, Sruthi Ramabadran, and Yusa Liu. 2021. "Alexa, What is That Sound?" A Video Analysis of Child-Agent Communication From Two Amazon Alexa Games. In *Proceedings of the 20th Annual ACM Interaction Design and Children Conference* (Athens, Greece) *(IDC '21)*. Association for Computing Machinery, New York, NY, USA, 513–520. https://doi.org/10.1145/3459990.3465195

[11] Teresa Flanagan, Gavin Wong, and Tamar Kushnir. 2023. The Minds of Machines: Children's Beliefs About the Experiences, Thoughts, and Morals of Familiar Interactive Technologies. *Developmental psychology* 59 (04 2023). https://doi.org/10.1037/dev0001524

[12] Michael A. Forrester. 2008. The Emergence of Self-Repair: A Case Study of One Child During the Early Preschool Years. *Research on Language and Social Interaction* 41 (2008), 128 – 99. https://api.semanticscholar.org/CorpusID:144424484

[13] Anja Gampe, Katharina Zahner-Ritter, Joanna Joys Müller, and Sarah Rebecca Schmid. 2023. How children speak with their voice assistant Sila depends on what they think about her. *Computers in Human Behavior* 143 (2023), 107693. https://doi.org/10.1016/j.chb.2023.107693

[14] Nicole Gardner-Neblett. 2022. What Predicts Oral Narrative Competence Among African American Children? Exploring the Role of Linguistic and Cognitive Skills. *Journal of Speech, Language, and Hearing Research* 65, 8 (2022), 2931–2947. https://doi.org/10.1044/2022_JSLHR-22-00002 Epub 2022 Aug 1.

[15] Nicole Gardner-Neblett. 2024. Becoming fictional storytellers: African American children's oral narrative development in early elementary school. *Child Development* (2024). https://doi.org/10.1111/cdev.14075

[16] Heather M. Gray, Kurt Gray, and Daniel M. Wegner. 2007. Dimensions of Mind Perception. *Science* 315, 5812 (2007), 619–619. https://doi.org/10.1126/science.1134475

[17] Alexis Hiniker, Amelia Wang, Jonathan Tran, Mingrui Ray Zhang, Jenny Radesky, Kiley Sobel, and Sungsoo Ray Hong. 2021. Can Conversational Agents Change the Way Children Talk to People?. In *Proceedings of the 20th Annual ACM Interaction Design and Children Conference* (Athens, Greece) *(IDC '21)*. Association for Computing Machinery, New York, NY, USA, 338–349. https://doi.org/10.1145/3459990.3460695

[18] D. L. Hughes, L. McGillivray, and M. Schmidek. 1997. *Guide to Narrative Language: Procedures for Assessment*. Thinking Publications.

[19] Hannah M. Julien, Lizbeth H. Finestack, and Joe Reichle. 2019. Requests for Communication Repair Produced by Typically Developing Preschool-Age Children. *Journal of Speech, Language, and Hearing Research* 62, 6 (2019), 1823–1838. https://doi.org/10.1044/2019_JSLHR-L-18-0402

[20] Joanne A. Konefal and Joann Fokes. 1984. Linguistic analysis of children's conversational repairs. *Journal of Psycholinguistic Research* 13, 1 (1984), 1–11. https://doi.org/10.1007/BF01067899

[21] Jacqueline M. Kory-Westlund and Cynthia Breazeal. 2019. Assessing Children's Perceptions and Acceptance of a Social Robot. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children* (Boise, ID, USA) *(IDC '19)*. Association for Computing Machinery, New York, NY, USA, 38–50. https://doi.org/10.1145/3311927.3323143

[22] Silvia Lovato and Anne Marie Piper. 2015. "Siri, is this you?": Understanding young children's interactions with voice input systems. In *Proceedings of the 14th International Conference on Interaction Design and Children* (Boston, Massachusetts) *(IDC '15)*. Association for Computing Machinery, New York, NY, USA, 335–338. https://doi.org/10.1145/2771839.2771910

[23] Lina Mavrina, Jessica Szczuka, Clara Strathmann, Lisa Bohnenkamp, Nicole Krämer, and Stefan Kopp. 2022. "Alexa, You're Really Stupid": A Longitudinal Field Study on Communication Breakdowns Between Family Members and a Voice Assistant. *Frontiers in Computer Science* 4 (01 2022). https://doi.org/10.3389/fcomp.2022.791704

[24] Ivonne Monarca, Franceli L. Cibrian, Angel Mendoza, Gillian Hayes, and Monica Tentori. 2020. Why doesn't the conversational agent understand me? a language analysis of children speech. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers* (Virtual Event, Mexico) *(UbiComp/ISWC '20 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 90–93. https://doi.org/10.1145/3410530.3414401

[25] Tova Most. 2002. The Use of Repair Strategies by Children With and Without Hearing Impairment. *Language, Speech, and Hearing Services in Schools* 33, 2 (2002), 112–123. https://doi.org/10.1044/0161-1461(2002/009)

[26] Elizabeth S. Nilsen and Susan A. Graham. 2009. The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology* 58, 2 (2009), 220–249. https://doi.org/10.1016/j.cogpsych.2008.07.002

[27] Cansu Oranç and Aylin C. Küntay. 2020. Children's perception of social robots as a source of information across different domains of knowledge. *Cognitive Development* 54 (2020), 100875. https://doi.org/10.1016/j.cogdev.2020.100875

[28] Barbara J O'Keefe and Jesse G Delia. 2018. Psychological and interactional dimensions of communicative development. In *Recent advances in language, communication, and social psychology*. Routledge, 41–85.

[29] Rachel L. Severson and Kristi M. Lemm. 2016. Kids see human too: Adapting an individual differences measure of anthropomorphism for a child sample. *Journal of Cognition and Development* 17, 1 (2016), 122–141. https://doi.org/10.1080/15248372.2014.989445

[30] Nancy Stein and Christine Glenn. 1979. An Analysis of Story Comprehension in Elementary School Children. *New Directions in Discourse Processing* 2 (01 1979).

[31] Mike van Duuren and Michael Scaife. 1996. "Because a robot's brain hasn't got a brain, it just controls itself" — Children's attributions of brain related behaviour to intelligent artefacts. *European Journal of Psychology of Education* 11 (12 1996), 365–376. https://doi.org/10.1007/BF03173278

[32] Randi Williams, Hae Won Park, and Cynthia Breazeal. 2019. A is for Artificial Intelligence: The Impact of Artificial Intelligence Activities on Young Children's Perceptions of Robots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3290605.3300677

[33] Ying Xu. 2023. Talking with machines: Can conversational technologies serve as children's social partners? *Child Development Perspectives* 17, 1 (2023), 53–58. https://doi.org/10.1111/cdep.12475

[34] 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* (Chicago, IL, USA) *(IDC '23)*. Association for Computing Machinery, New York, NY, USA, 160–172. https://doi.org/10.1145/3585088.3589354

[35] 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.

[36] Ying Xu and Mark Warschauer. 2020. Exploring young children's engagement in joint reading with a conversational agent. In *Proceedings of the Interaction Design and Children Conference* (London, UK) *(IDC '20)*. Association for Computing Machinery, New York, NY, USA, 216–228. https://doi.org/10.1145/3392063.3394417

[37] Ying Xu and Mark Warschauer. 2020. What Are You Talking To?: Understanding Children's Perceptions of Conversational Agents. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376416

[38] Kristine Yont, Lynne Hewitt, and Adele Miccio. 2000. A Coding System for Describing Conversational Breakdowns in Preschool Children. *American Journal of Speech-Language Pathology* 9 (11 2000), 300. https://doi.org/10.1044/1058-0360.0904.300

[39] Chao Zhang, Xuechen Liu, Katherine Ziska, Soobin Jeon, Chi-Lin Yu, and Ying Xu. 2024. Mathemyths: Leveraging Large Language Models to Teach Mathematical Language through Child-AI Co-Creative Storytelling. *Arxiv* abs/2402.01927 (2024). https://doi.org/10.48550/arXiv.2402.01927

[40] Zheng Zhang, Ying Xu, Yanhao Wang, Bingsheng Yao, Daniel Ritchie, Tongshuang Wu, Mo Yu, Dakuo Wang, and Toby Jia-Jun Li. 2022. StoryBuddy: A Human-AI Collaborative Chatbot for Parent-Child Interactive Storytelling with Flexible Parental Involvement. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 218, 21 pages. https://doi.org/10.1145/3491102.3517479

A PARTICIPANT INFORMATION

See Table 3 in the next page.

B CODEBOOK FOR COMMUNICATION BREAKDOWNS AND REPAIR STRATEGIES

B.1 Communication Breakdowns Types Coding Protocol

See Table 4 in the next page.

B.2 Repair Strategies Types Coding Protocol

See Table 5 in the next page.

Table 3: Participants' information by study conditions

	AI	Human	Full	Difference
Female	53.85%	46.43%	62.50%	$\chi^2(1) = 0.992, p = .319$
Age				$\chi^2(4) = 3.065, p = .547$
4-5-year-olds	31.25%	46.15%	39.44%	
6-8-year-olds	68.75%	53.85%	60.56%	
Race/Ethnicity				$\chi^2(6) = 9.502, p = .147$
Black	6.25%	0.00%	2.86%	
Latinx/Hispanic	3.12%	7.89%	5.71%	
Asian	0.00%	15.79%	8.57%	
White	31.43%	30.00%	61.43%	
Mixed	18.75%	18.42%	18.57%	
Prefer not to say	3.12%	2.63%	2.86%	
Use of CA				$\chi^2(5) = 3.272, p = .658$
Daily	28.12%	18.42%	22.86%	
Weekly	21.15%	21.43%	21.43%	
Monthly	12.50%	10.53%	11.42%	
Less than once a month	18.75%	28.95%	24.29%	
Never	15.62%	23.68%	20.00%	
Total	32	39	71	

Table 4: Coding types for different communication breakdowns with descriptions and examples.

Code	Descriptions	Examples
Auditory aspect	Occurs due to the child or partner's utterances are not articulated clearly enough to be recognized by the other party.	AI: What do Logan and Erry will do to add more fun and learning to their adventure? Child: They will *(Pronunced as "Ay will")*. AI: A Will? That's very mysterious. I wonder who left this will and what secrets might it hold? Can you tell me more? Child: Yeah...
Verbal aspect	Information provided to the child or partner is too complex for the context, leading to confusion; this may involve the use of advanced vocabulary and concepts, incorrect grammar, or complex sentence structures.	Human: So... The cherries and feathers are the same amount. They pour this into the pot. And the sum of all the cherries... Child: And all tiny arrows. Have big arrows on my computer... Like I have it this big... Human: That's pretty big? *(Looked confused)* Child: The whole right... this big. Human: I don't know.
Contextual aspect	Applies when the child fails to provide enough context for the partner to understand the message. This could be due to assuming shared knowledge that the listener does not have, or failing to establish a common frame of reference in the conversation.	Human: So... Do you want it to be an animal or like a human? Child: No, I want it to be a bus (Referred to a cartoon about goldfish bus). Human: A bus? *(Looked confused)* Child: No, it's like a goldfish. Human: Oh, a goldfish.
Engagement aspect	Identifies moments when communication is hindered by external distractions affecting the child. Distractions may include environmental noises, interruptions by others, or the child's attention shifting away from the interaction.	Human: So how do you think that they're going to conquer the waves in the ocean? Child: *(Playing with the plush toy)* Human: What do you think? do you have any ideas? Human: What do you think? do you have any ideas? Child: *(Playing with the plush toy)*

Table 5: Coding types for different repair strategies with descriptions and examples.

Code	Descriptions	Examples
Repetition	Is assigned when the strategy used to address a communication breakdown is repeating the original message. Repetition can be exact or with slight modifications for emphasis.	AI:That's interesting. What's the character's name? Child: Redco *(Pronounced as "Greco")*. AI: Great! Where will our story happen? Child: It'll happen at the store of a dragon. AI: Okay. We'll dive into an incredible adventure with a zebra named Greco L. Heaven in the story of a dragon. Ready to start the story? Let's go. Once upon a time, in the heart of the Serengeti lived a zebra named Greco L.Heaven... Child: Redco *(Interjected)*.
Correction	Applies when a child corrects the AI or human partner by clearly stating what they actually meant. They might use phrases like "No, it's not..." or "I didn't say that..." to point out the mistake and provide the correct information.	AI: Your introduction of a beetle bug in a book adds a fun layer to the story. Great job! They found a beetle bug, which was equal in size to the book they carried... How do you think they will use the beetle bug that is equal to the size of their book, and what they might subtract from their day to read it. Child: I didn't say beetle bug in in a book *(Had said "Read a book" before)*.
Request for clarification	Applies when a child indicates confusion or a lack of understanding. This could be through verbal cues, such as saying "I don't understand," "I don't know what it meant," "Huh?", "What?" or non-verbal cues like a puzzled expression.	AI: How do you think they will estimate the time state by subtracting the obstacle of the bear's territory? Child: I cannot understand the question.
Accommodation	Demonstrates instances where an AI misinterprets a child's input, yet the child adjusts their responses to match the AI's feedback. The child does this even though the direction of the conversation deviates from their original intent.	AI: What do you think they will subtract from their belongings to make their journey more efficient? Child: Uh... Clinton [anonymized child's name] steals *(laughed)*. AI: Oh no! Plankton is playing tricks again. What does he steal this time? And how does he do it? Child: He said Plankton *(Laughed and said to the researcher)*. Well… they're having a cook over a fire, he comes up and sneaks one up the front and two of the back.