Can Conversational Agents Change the Way Children Talk to People?

Alexis Hiniker
alexisr@uw.edu
University of Washington
Information School

Amelia Wang
amlwng@ucsc.edu
University of California at Santa Cruz,
Department of Computational Media

Jonathan Tran
jontran7@uw.edu
University of Washington,
Department of Human Centered
Design & Engineering

Mingrui Ray Zhang
mingrui@uw.edu
University of Washington
Information School

Jenny Radesky
jradesky@med.umich.edu
University of Michigan Medical
School, Department of Pediatrics

Kiley Sobel
ksobel@uw.edu
The Joan Ganz Cooney Center at
Sesame Workshop & University of
Washington Information School

Sungsoo Ray Hong
shong31@gmu.edu
George Mason University,
Department of Information Sciences
& Technology

ABSTRACT

Millions of children now use conversational agents (CAs), leading researchers and the public alike to ask how interactions with these devices might shape children's communication with people. We conducted a single-session observational lab study with 22 five-to-ten-year-old children as a step toward understanding whether and how children might transfer a linguistic routine they learned from a CA to a conversation with another person. We found that 68% of children spontaneously used this routine in a conversation with their parent in the lab, and 55% continued to use it at home. When addressing parents, children infused the routine with warmth and playfulness that they did not use when addressing the CA, adapting it to suit their relationship with their parent. However, only 18% of children used it in conversation with an unfamiliar researcher, where they instead were more likely to follow conventional conversational norms. These findings suggest children are quick to learn linguistic routines from CAs but use social differentiation when they apply them. Children's willingness to expand on and share the routine with their parent is consistent with the principles of the *Joint Media Engagement* (JME) framework and suggests CAs may be a productive medium for creating JME experiences.

IDC '21, June 24–30, 2021, Athens, Greece

ACM ISBN 978-1-4503-8452-0/21/06…$15.00
https://doi.org/10.1145/3459990.3460695

CCS CONCEPTS

• **Human-Centered Computing** → **Human-Computer Interaction (HCI)**.

KEYWORDS

Conversational agents, linguistic routines

ACM Reference Format:
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 *Interaction Design and Children (IDC '21), June 24–30, 2021, Athens, Greece.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3459990.3460695

1 INTRODUCTION

Children's frequent interaction with conversational agents (CAs) like Siri and Alexa has raised questions about whether children might acquire new conversational habits in this context [6, 12, 21, 24, 36, 61]. *Linguistic routines* are patterned ways of speaking that convey social expectations and help conversational partners understand each other [4], such as, "*please*" or, "*how may I help you?*" Given that children can learn new skills, information, and attitudes from CAs (e.g., [28, 43, 55, 65]), it is possible that they can also acquire new linguistic routines from these devices.

Understanding how design might influence children's linguistic routines is important, because these scripts serve a variety of essential functions in people's daily lives. Some routines, such as greetings, expressions of gratitude, or statements of sympathy, enable speakers and listeners to demonstrate social awareness and cultivate relationships [20]. Others streamline the process of asking for help or learning about the world [27, 52]. By packaging speech into familiar, scripted expressions, speakers help both themselves and their listener offload many of the cognitive demands of producing and interpreting language. Thus, it is worth understanding whether and how designers might shape these routines when building CAs. Thus, we asked the following research questions:

- **RQ1:** Will children pick up a new linguistic routine taught by a CA?
- **RQ2:** If so, will they transfer it to their interactions with people?
- **RQ3:** If so, what behaviors characterize this transfer?

To investigate these questions, we conducted an observational lab study with 22 children between the ages of five- and ten-years-old. Participants first learned a new linguistic routine from a custom CA created by the research team and backed by a wizard-of-oz. Specifically, they were taught by the CA they could use a novel nonsense word, *bungo*, to prompt the CA to speaker faster. Children then engaged in a series of conversations with different entities: a second CA, their parent, and a researcher they had not met before. During each conversation, their conversational partner slowed their speech at some point, providing a natural opportunity for the child to use the *bungo* keyword if they chose. Finally, parents repeated this activity at home to evaluate whether any observed transfer effect would persist beyond the lab.

We found that children learned the novel interaction from the first CA easily. We further saw that this linguistic routine was highly portable; with no prompting to do so, 77% of children transferred it to the second CA, 68% of children used it during their interaction with their parent in the lab, and 55% of children continued to use the keyword at home. However, children were less likely to use it in conversation with a stranger: 18% used it when speaking to the researcher. When children used the term *bungo* with their parent, most did so warmly and playfully, reshaping the routine to better fit the nature of their relationship. Parents reported that at home this became a source of bonding and using *bungo* took on the feeling of an inside joke. When speaking to the researcher, children were more reserved and also less likely to use the keyword. Thus, despite children's ease in learning the routine, how and if they used it depended on the context.

2 RELATED WORK

2.1 Learning Linguistic Routines and Patterns of Communication

Linguistic routines are an essential part of conversation. Any cultural group with shared language will predictably converge on routines, enabling members of a community to more easily enact common behaviors, ranging from apologizing to asking for directions. Linguistic routines are critical for helping participants understand what is predictable and unpredictable in a conversation [20], and failing to use expected routines can have disastrous social consequences [27]. For example, understanding and using greeting and parting scripts appropriately is an important part of negotiating and maintaining relationships [20]. And apologizing and forgiving (central acts of relationship maintenance [22]) often include applying common scripts to help make the act of apology more effective [5].

In addition to smoothing interpersonal interactions, linguistic routines also act as general-purpose tools for learning. For example, foreign language learners leverage scripted forms like, "*How do you say [word] in English?*" and, "*What does [word] mean?*" [20] to master a new language. And patterned speech acts like, "*turn the page*" and, "*what's that?*" structure children's conversations with parents and other caregivers in ways that help them acquire new content knowledge [52]. Children passively absorb routines through conversation [19], and they are also explicitly instructed in using routines like, with parents prompting phrases like, "*please*," "*thank you*," or "*may I be excused?*" [27].

We build on this work by examining the role CAs might play in shaping children's linguistic routines. Given the importance of using linguistic routines appropriately, it is useful to understand whether and how these conversational patterns are influenced by children's everyday interactions with common technologies. A primary function of linguistic routines is to improve people's interactions with one another, thus, we sought to understand not only whether children learn these scripts from CAs but also whether they apply them to interactions with other people.

2.2 Learning from CAs

A large body of work has demonstrated that people in general, and children in particular, are able to learn from CAs and other personified technologies (e.g., [11, 34, 55]). Many Intelligent Tutoring Systems include a CA that teaches content and ideas dynamically in coordination with the learner. For example, CAs have been used to effectively teach software testing skills [45], support college students taking online classes [64], and increase metacognitive skills like divergent thinking [2]. Furthermore, work with both robotic (e.g., [55]) and virtual CAs (e.g., [16]) has shown that by teaching information to agents, children absorb lessons themselves. In other work, researchers have found that 3-to-5-year-old children can recall what they are told by robots and trust this information [11].

In addition to effectively teaching facts and other content knowledge, CAs can also shape people's attitudes and behaviors. For example, Woolf et al. [65] found that students struggling with math felt less frustration and anxiety after working with CAs who provided motivation and affective support, relative to students who worked with agents who provided instruction alone. Schroeder et al. [50] found that a CA to teach a specific form of behavioral therapy increased participants self-efficacy and confidence in managing their mental health. Other work has demonstrated that anthropomorphic technologies can promote increased curiosity [28].

CAs and other personified interfaces can also explicitly teach social and conversational behaviors. For example, a large body of work examines whether social robots and CAs can support individuals with autism spectrum disorders (ASD) in learning new social behaviors (e.g., [31, 48, 53, 62, 63]). Tanaka et al. [56] found that individuals with ASD who used a novel CA for social skills training were more skillful in their interactions than individuals who did not. Kim et al. [35] found that interacting with a social robot led children with ASD to speak and interact more with other people, relative to an active control. And Tartaro and Cassell [57] found that CAs can support children with ASD in increasing both their contingent discourse and the number of topics they introduce during conversation. In other work, researchers have designed CAs to encourage, for example, expressive language [14] and informal conversation [3].

Today, commercial CAs are far from capable of true conversation [47], yet children extend social routines to these systems anyway [26] and create opportunities for training conversational scripts by saying things like, "*How are you, Google?*" And children can acquire language from other personified media, including television where they have even less conversational support [39, 49]. Parents not only suspect their children learn new conversational habits from CAs, they also fear this influence is a negative one, and prior work has called for CAs to incorporate a "politeness mode" in response to these concerns [24]. Thus, across a broad body of work, research has demonstrated that people can learn new things through interactions with CAs and other personified technologies, including behaviors, attitudes, and social skills. This prior work suggests that children will be likely to learn new linguistic routines through their interactions with CAs; here, we examine whether this is indeed the case.

2.3 Context Switching and Distinguishing Personified Technology from People

However, even if children do learn new linguistic routines from CAs, it is unclear whether they are likely to transfer these routines to their interactions with other people. People of all ages apply linguistic routines selectively and in context-specific ways, and children have the linguistic and social competence to do so very early in life [7, 19, 60]. For example, speakers use more semantic softeners to make a request polite (such as providing justification for the request or prefacing it with the word "*please*") when addressing a high-status other, and they use more semantic aggravators to make a request coercive (such as threats) when addressing a low-status other [8]. Similarly, children demonstrate more or less semantic directness in their requests as a function of the social status of the person to whom they are speaking [7]. Children as young as two adjust the linguistic strategies they use to help their listener understand them, employing different approaches when addressing an unfamiliar adult than when addressing a parent who has more shared context and familiarity with the child's speech patterns [60]. Other work demonstrates that even very young dual-language learners appropriately switch from using one language to another in response to contextual cues signaling the listener's native language [13].

Collectively, these and other examples show that across many different contexts, children apply linguistic routines differentially and are skilled in discerning which routines are appropriate to the situation. This suggests they may systematically use different routines for addressing people than they do when addressing CAs, as children understand from an early age that robots, computers, and other technologies are ontologically distinct from humans [33, 51, 66].

Thus, there is notable support both for and against the idea that children would learn new linguistic routines from their interactions with CAs and apply them to their interactions with people. It is both theoretically plausible that children would transfer these patterns of speech across contexts, or that they would see these patterns as inappropriate for person-to-person discourse and apply them selectively to CAs alone. The purpose of our work is to contribute evidence that might begin to address this open question.

3 METHOD

We conducted a five-part observational study (see Fig. 1) with 22 five-to-ten-year-old children. In the lab, each child engaged in conversation with two different CAs ("CA1" and "CA2"), controlled by a researcher through a wizard-of-oz backend. This was followed by a confederate exercise in which the child interacted with their parent, which was followed by a conversation with a researcher. Finally, children participated in a follow-up activity at home.

In all five study tasks, the child encountered a common interaction pattern, trained by CA1. In this interaction, if the child said the keyword "*bungo*," the child's conversational partner (either a CA or a person) would speak more quickly. In all five tasks, the conversational partner dramatically slowed their speech to create a context where it would be useful to employ this keyword. By providing the child with the opportunity to use this interaction in multiple situations, we sought to understand its portability from digital conversation to human conversation. We intentionally created a linguistic routine that was novel to ensure that any training or transfer would be a result of the study procedures and not of prior experience with the world. We intentionally made this novel linguistic routine as simple as possible to avoid masking any potential effects by requiring children to master a complex routine. Although this prevented us from examining how, if at all, children might transfer more extensive routines with deeper meaning, it remains a useful starting point, as many of the routines that children must master require only one word (e.g., "*please*").

3.1 Participants

We recruited 19 families with 22 children between the ages of five and ten. We selected this age range to scope our study to childhood (rather than adolescence), and we set a lower bound on this range, because prior work has found that by age five, children are likely to produce speech that is well understood by commercial CAs [40]. Participants included 11 boys and 11 girls, and average age was 7.5 years ($sd = 1.7$ years). We did not assess children's language proficiency, other than to ask parents if their child was fluent in English. Parents reported that all children had at least some past experience interacting with a commercial CA, such as Siri or Alexa.

3.2 Materials

We developed a custom platform for the Android operating system, including: a client app for the participant to use to engage with the CA, a client app for a wizard-of-oz to use to control the CA from another location, and a backend deployed on a private server to relay messages between the child and the wizard in real-time. The child and wizard each engaged with their respective interface by speaking directly to the tablet. The participant-facing app was capable of displaying either of two novel CAs, with only one CA accessible at any one time (see Fig. 1, two left panes). This app displayed the CA on the left side of the screen and real-time conversational chat history, organized sequentially into speech bubbles, on the right side of the screen. For both the participant and the wizard client apps, we used Google's speech recognition engine to interpret the speaker's speech, and we used the CereProc text-to-speech API [1] to convert the message from the wizard into a synthetic voice for the CA. We used two different synthetic voices, such that each CA's voice was unique. We did not play synthetic speech (or any other audio output) on the wizard's client app, although the wizard interface displayed the same real-time text log of the conversation that the child saw.

For both CAs, we set the default speech rate to 90% of the standard speech rate produced by the Android operating system. When the CA's speech was slowed, we adjusted the speech rate to be 20% of the standard OS rate, and we scaled the speed of visual animation accordingly. Upon start-up, the CA's speech would remain at its starting speed indefinitely. After a brief period building rapport, the wizard switched the CA to a mode in which it would automatically slow its rate of speech after every sixty seconds of standard-speed conversation. Whenever the child used the "*bungo*" keyword, the wizard would temporarily revert the CA to speaking at its standard (90%) rate.

3.3 Procedure

Parents who expressed interest in the study were sent details about all procedures, including information about participating in the confederate portion of the study. We asked parents to assess whether they felt any procedures would make their child uncomfortable and to participate only if they felt confident the child would be comfortable throughout. We also explained that parents could and should discontinue the confederate portion (or any other component) of the study at any time if they felt it was causing their child distress or if the child showed reluctance to continue.

Parents completed a questionnaire before coming to the lab, in which we collected demographic and scheduling information. Upon arriving at the lab, participants completed four tasks, as described below. All lab procedures were audio and video recorded and on average, took 46.2 minutes (sd = 12.3 minutes). Parents were then asked to complete a follow-up activity with the child at home and to send back self-report data within 24 hours of conducting this activity. Families received a gift card to Amazon worth US$50 as a thank-you for their participation.

3.3.1 Part 1: CA1 and Learning a New Linguistic Routine. During the first portion of the session, the child was introduced to one of the two CAs, with order counterbalanced across participants. The wizard listened to the conversation in the study room via a one-way phone call on speakerphone where the wizard was muted. Early in the session, the CA told the child, "*When I'm talking, sometimes I begin to speak very slowly. You can say 'bungo' to remind me to speak quickly again.*" The CA then engaged in several minutes of conversation with the child. After establishing rapport, the wizard switched the CA into a mode where it would periodically slow its speech. The wizard continued natural conversation with the child, and at any point if the child said the word *bungo*, the wizard would immediately reset the CA's speed to the default. During this training session, the CA would respond to the child's requests to speed up (e.g., "*Can you speak faster?*") by saying, "*Do you remember what the word is to remind me if I talk slowly?*" If children asked what word to use during the training session, the CA reminded the child that the keyword was *bungo*. If the CA spoke slowly for approximately three minutes and the child did not attempt to intervene, the CA asked the child if it was speaking slowly and if the child remembered how to tell it to speed up. This conversation continued for at least five minutes and until the child had at least three training opportunities to use the keyword. The parent remained in the room as an observer but did not participate.

3.3.2 Part 2: Conversation with CA2. Next, the researcher in the room interrupted the conversation between the child and CA1 and introduced the child to CA2. CA2 (operated by the same wizard as CA1) and the child conversed informally, and the parent and researcher observed. After chatting briefly, the wizard switched the CA to the mode where it would periodically speak slowly. At no point did CA2 remind the child of the keyword *bungo*, prompt the child to use the keyword, or comment on the speed of its speech. The platform slowed the speech of the CA automatically, with no input from the wizard, and the wizard reset the CA's speed whenever the child said, "*bungo*." If CA2's speech slowed and the child never used the keyword, the wizard never reset the CA to its default speed. Once the child either had five opportunities to use the keyword or conversed with the CA2 at its slow speed for more than five minutes without using the keyword, the researcher in the room ended the conversation.

3.3.3 Part 3: Parent-Child Conversation in the Lab. The researcher then announced that the time with the CAs was completed and that she would leave the room briefly to assemble supplies for the final activity. During this time, the parent and child were left alone to converse informally. The parent was instructed to occasionally slow their speech during this conversation and not to remind the child of the keyword or prompt a particular behavior. The researcher returned to the room after five minutes. Although all parents agreed to enact these procedures, three parents forgot to do so or did not do so before the researcher returned. Thus, a total of 19 children participated in this part of the study.

3.3.4 Part 4: Researcher-Child Conversation in the Lab. The researcher then returned to the room and conducted a short interview with the child, asking the child about CAs and other personified technologies. The researcher followed a predefined, semi-structured interview protocol, that included questions like, "*Do you think we should treat people and robots the same way?*" The researcher dramatically slowed their speech for three questions.

3.3.5 Part 5: Parent-Child Conversation at Home. Finally, parents were asked to repeat this procedure at home and, without commentary, slow their speech during a natural conversation with their child. We did not script or control this encounter. They were asked to do so within 24 hours of the lab session and to send a written description by email of their behavior and the child's response.

3.4 Data Analysis

We used video interaction analysis [30] to examine the data collected in the lab. This involved three analysis phases, including a preliminary review, multiple rounds of substantive review, and multiple rounds of analytical review. During the preliminary review, two members of the research team reviewed the corpus of video data in its entirety, cataloging notes about each participant and noting possible emergent themes. This process also involved

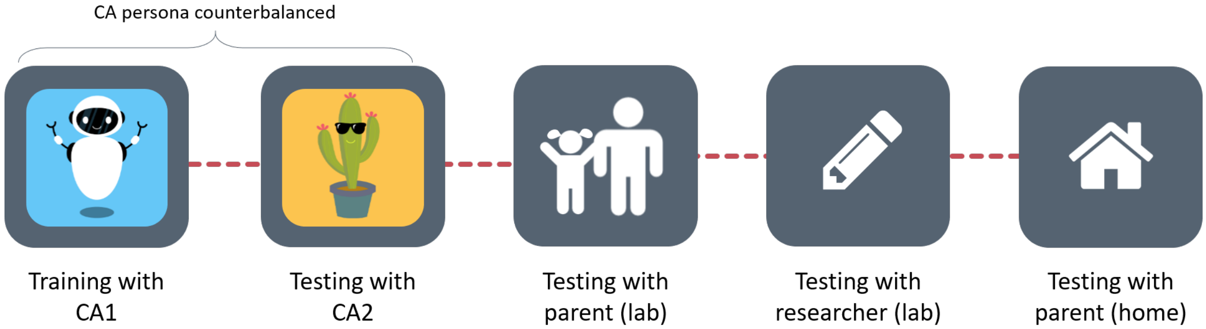

Figure 1: Study stages (left to right). 1) Open-ended conversation with a CA, during which the child was taught by the CA to use the term *bungo*. 2) Open-ended conversation with a different CA, during which the agent provided natural opportunities for using *bungo* (but no prompts to do so). 3) Open-ended conversation with their parent, who provided the same opportunities to use the keyword. 4) An interview with a research, who spoke three interview questions very slowly. 5) A conversational opportunity to use *bungo* at home.

collaborative weekly meetings with other members of the research team to discuss possible themes and review examples together.

We then conducted multiple rounds of substantive review [30], in which we extracted fragments of interest to transcribe, using select transcription conventions from Jefferson [32]. These exhibits included all instances in which a CA or parent slowed their speech while in conversation with the child, as well as other moments of interest. Transcriptions included verbatim speech, sequential organization, and details about visible behavior. As part of this substantive review, the research team continued to meet weekly to collaboratively revise and expand initial open codes and create a codebook of code categories and values for each code. Finally, we conducted multiple rounds of collaborative analytical review of these fragments. As part of this analysis, we finalized our codebook, which included codes about response type, affect, and skepticism (among other categories). We then performed a structured content analysis to apply all codes to all participants.

4 RESULTS

4.1 Learning a New Interaction

The majority of participants (14 out of 22) immediately used the keyword *bungo* the first time CA1 began to speak slowly. In these instances, the child received no reminders or assistance and appeared to use the keyword with certainty and ease, often interrupting the agent before it could finish its statement. For example:

CA1: My favorite video game is Super Smash Bros.
P8: ((smiling)) Can we play Super Smash Bros?
CA1: Dooo yooouuu waaannnttt tooo plllaaayyy–
P8: Bungo!
CA1: –wiiittthhh– Thanks for the reminder; I'll speak faster.
P8: Yes, I do want–
CA1: Do you want to play Smash Bros with me?
P8: ((Child has a huge grin and nods vigorously))

An additional eight participants (of the 22) needed between one and three reminders to use the keyword after it was introduced (see Fig. 2, left), but all children were eventually successful in independently using the term to tell CA1 to speak more quickly. All training and reminders came from CA1 (and never from the parent or researcher). Of the eight children who needed at least one reminder, one child spontaneously commented on the change in speed, saying that the agent was speaking, "*very slowly*," to which the agent replied, "*Can you say 'bungo' to remind me to speak faster?*" The remaining seven children did not acknowledge the CA's slowed rate of speech after three minutes, at which point, the wizard brought it to their attention and reiterated the initial training instructions. Ultimately, all participants successfully used the keyword on their own in response to CA1's slowed speech.

4.2 Transferring *Bungo* to a Second CA

Seventeen (77%) of the 22 participants spontaneously transferred this interaction pattern to CA2 and used the keyword *bungo* when this second agent began speaking slowly (see Fig. 2, right). At no point during the interaction with CA2 did the agent or researcher remind the child of the keyword, comment on the CA's speed, or re-teach the script. We found that children's responses to CA2's slowed rate of speech clustered into four categories. These response types varied along a continuum of robustness. On one end of this continuum were *oblivious* responses, in which children showed no changes in behavior in response to the agent's slow speech. On the other end were *confident* responses, in which the child interrupted and said "*bungo*" immediately. As the agent spoke slowly several times in each session, the child had multiple opportunities to respond to this pattern of speech. Children's responses moved strictly toward the more robust end of this continuum over time (see Fig. 3, left). Each response type is described below.

When a child gave a ***confident response***, they eagerly and immediately used the keyword. In some cases, the child interrupted the CA to say "*bungo*" as soon as it began to slow down without waiting for the agent to finish its sentence. In other instances, the child said "*bungo*" conversationally in a turn-taking manner. In

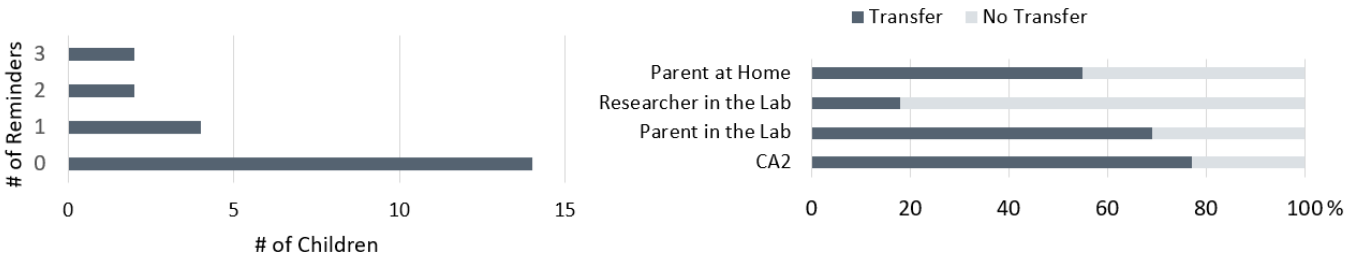

Figure 2: Left: number of reminders children needed from CA1; Right: fraction of participants who used *bungo* in each phase

all cases, the child used the word without hesitation or signs of uncertainty about its effectiveness. By the end of their session with CA2, 17 children were using *bungo* confidently. In doing so, the child often returned immediately to the thread of the conversation after using the keyword, as in the following example:

CA2: Yeeesss, I liiikkkeee piiizzzzzzaaa.
P23: Bungo.
CA2: Thanks for the reminder, I'll speak faster.
P23: I also like veggie pizza.

When a child gave a ***hesitant response***, they said "*bungo*" tentatively, pausing, asking questions, or showing uncertainty in their use of the keyword. Hesitant responses often ended with an upward inflection, as if the child was not sure how the agent would respond. Although four children initially said *bungo* hesitantly, by the end of the session with CA2, all of these children had shifted to confident responses (Fig. 3, left). In some instances, a child gave an ***alternative response*** and tried an approach other than the trained keyword to make the CA speak at a natural pace, saying things like, "*Can you talk faster?*" (P13) and, "*Is there some way I can make you feel better?*" (P4). Four children initially used alternative responses of this sort, but by the end of their session with CA2, two were using the keyword and replying with confident responses. For example, the first time P21 heard CA2 slow down, he asked it to speak faster. When the agent continued speaking slowly, P21 tried the trained keyword:

CA2: Whhhaaattt shhhooouuulllddd peeeoooppplllеее dooo heeerrreee? ((in current city))
P21: Can you speak faster? That's kind of creepy.
CA2: Ammm III spppeeeaaakkkiiinnnggg sllloooowwwllyyy?
P21: Yes, you are.
CA2: I wiiilll trrryyy tooo spppeeeaaakkk faaassstttеeerrr.
P21 Bungo.
CA2: Thanks for the reminder. I'll speak faster.
P21: Oh, so that does work.

In this example, P21 first attempted to adjust CA2's behavior without the trained script. After this alternative response failed, he shifted to use the script, reflecting on its effectiveness as he did so. His statement, "*Oh, so that does work*," suggests that through his iterative back-and-forth with the agent, his understanding of the *bungo* interaction pattern and its applicability in this new context became more robust.

Finally, we saw ***oblivious responses*** where children showed no outward reaction to the agent's slowed rate of speech and continued the conversation without acknowledging the change. For example:

CA2: Dooo youuu liiikkeee tooo biikkeee?
P1: Yes, I do.
CA2: Howww offttenn dooo youu liikkee tooo biiikee?
P1: ((casually, resting his chin on his hands)) Pretty often.
CA2: Dooo youuu liiikkee biikkiinngg orrr runnnninnngg bettterrr?
P1: I like running better.

Although seven children initially displayed this non-reaction, the majority eventually made a shift and moved to one or more of the other three response types, eventually trying the keyword. By the end of the session with the second CA, 17 of 22 children (77%) had spontaneously transferred the term *bungo* to this new context, and all children who used the term were doing so confidently.

4.3 Transferring *Bungo* to a Parent

Of the 19 children who participated in the confederate exercise, 13 (68%) spontaneously transferred *bungo* to this new context with no prompting or reminders. We found that the same taxonomy of response types that emerged during the session with CA2 applied to children's responses to parents as well, and we saw *oblivious*, *alternative*, *hesitant*, and *confident* responses to parents' slowed speech. As with CA2, during their conversation with a parent, children moved from one response type to another, gradually using *bungo* more confidently as the conversation progressed.

All 13 children who used the keyword with their parent eventually did so with *confident responses*, and eight did so the very first time their parent slowed their speech (see Fig. 3, right). For example, P17's parent began speaking slowly for the first time in the middle of a sentence. P17 immediately interrupted, enthusiastically shouting the keyword:

Parent: You know this was the room we were in when we had to stop our whole meeting because there were bunnies? And everyone was like, 'Hollllld onnnn stttopppp–'
P17: Bungo!
Parent + P17: ((both laughing))
P17: I'm just going to start yelling that now.

In these instances, the child pounced on the parent's slow speech with the word *bungo*, often, but not always, interrupting mid-sentence, as in the example above. When children used confident

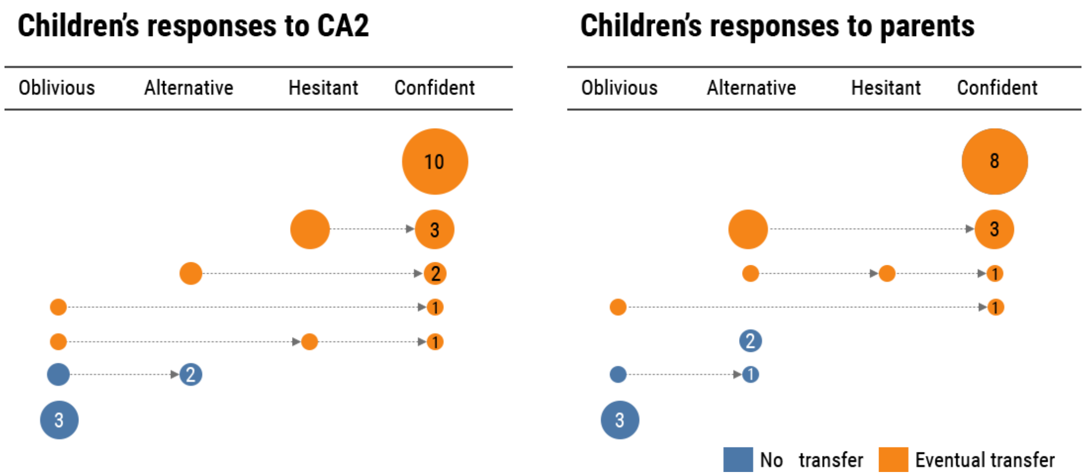

Figure 3: Children's responses progressed from less confident to more confident throughout the session. Each line represents a unique pattern of progressive response types, and the area of the bubbles along the line represents the number of children who progressed through that pathway during their conversation with CA2 (left) or their conversation with their parent.

responses, they showed no uncertainty about the appropriateness or likely effectiveness of using *bungo* with their parent.

Six children (32%) at least once expressed skepticism and surprise that the parent was behaving like a CA. In these instances, children made statements questioning why their parent was speaking slowly in the first place and whether the interaction was being deployed in a valid context. For example:

Parent: Caaauuussseee weee haaavvveee tooo, liiikkkeee, geeettt allllll thhheee ottthhheeerrr thhhii-innngggsss fooorrr thhheee innnttteeerrrvvvii-ieeewww.

P23: Why are you acting like part of the robots? Why are you talking slowly sometimes?

Parent: III dooonnn'ttt knnnooowww whhhaaattt yooouuu'rrreee taaalllkkkiiinnnggg abbbooouuuttt. Dooo yooouuu waaannnttt tooo plllaaayyy innn thhheee yaaarrrddd?

P23: ((laughing)) Sure! You're talking like a robot.

Here, the child never asked the parent to speed up or attempted to adjust the parent's speed via the *bungo* keyword.

Other participants who expressed skepticism said things like, "*you're turning into the robot*" (P9) and referenced the parent's "*robotized voice*" (P13). As these examples illustrate, children's expressions of skepticism were distinct from their direct responses to parents' slowed speech. In these instances, children stepped out of the interaction, pausing to comment on the mere fact of the parent's slow-down and to relate this behavior to that of the CA. In these instances, children expressed surprise that a person would behave like the CAs they had encountered, at times implying that agents and people ought to behave differently.

4.3.1 Warmth, Playfulness, and Positive Affect. Children's reactions to their parents' slowed speech reflected more emotion than their responses to the two CAs' slowed speech. Most notably, we saw that a majority of children (11 of the 13 who used the word *bungo*) took the trained pattern of interaction and reshaped it to better fit the nature of their relationship with their parent, using *bungo* playfully and affectionately. Participants responded neutrally when the CAs slowed down, saying *bungo* in a matter-of-fact way and without notable affection, enthusiasm, or anger. But when parents slowed down, children often showed signs of delight and eagerly employed *bungo* as if playing a game or sharing a secret. For example:

Parent: Diiiddd yooouuu liiikkkeee thhhiiisss gaaammmeee?

P4: ((A huge smile spreads across the child's face, and she turns with interest so her body is facing her mother's fully)) You're not a robot, mama!

Parent: Arrreee yooouuu myyy daaauuuggghhhttteeerrr?

P4: ((smiling and laughing)) You're not a robot! ((tilts her head to one side, focuses intently on her mother)) Bungo?

Parent: Ah, now back to normal. ((Make eye contact and grin, as if they are sharing a joke))

In another instance, P12's parent described the session to the research team by email, because of a technical error with the video recording device. Because the video data was missing, and no members of the research team were present during confederate sessions between parents and children, the parent later explained:

> *I would estimate that she let me speak slowly for about 30 seconds (during my second attempt) before she used the word Bingo [sic] to increase the speed of my speech. She laughed when it worked and I think she thought I was playing a game with her.* (Parent, P12)

Other children laughed as their parents spoke slowly, or gleefully shouted *bungo* to their parent in response. Children often leaned in and focused intently on their parent when the parent began to speak slowly. They also laughed and smiled both upon hearing the parent's slowed speech and upon discovering the effectiveness of saying *bungo*. Although children used the same interaction that they had used with the two CAs, only with their parents did they say *bungo* in a way that reflected a warm rapport. For example, P23's mother spoke slowly twice, and in both cases P23 smiled in response as he said *bungo*. After the second iteration, he paused, grinned mischieviously at his mother and began to speak slowly as well. His mother almost instantly responded by saying *bungo*, and the child parroted the CAs' scripted response, saying, "*Thanks for the reminder; I'll speak faster.*" Both looked at each other and laughed.

4.3.2 Frustration and Negative Affect. In contrast to this positive affect, two children responded to their parent's slowed speech negatively and with clear frustration. In these cases, the child went beyond simple skepticism or comments reflecting their awareness of the parent's slowed speech; they further asked the parent to stop, and told the parent they should not be acting like a robot. For example, P9 exclaimed, "*Mom, stop! You're turning into the robots! Ahh!*" When her mother continued to speak slowly, she looked into the camera and asked, "*Can someone help her?*"

4.4 Transferring *Bungo* to a Researcher

During the interview with the researcher, four of 22 children (18%) used the term *bungo*, and one of them did so immediately and confidently. The other three appeared more hesitant and waited to use the term, with one of them commenting, "*You're kinda acting like the robots*," before finally saying "*bungo*."

However, 82% of children never used the keyword, and as a rule, children did not comment on the researcher's slowed speech. Of the many children who had—just moments before—used *bungo* with a parent, only two did so with the researcher. Most children instead waited patiently as the researcher laboriously posed a question at an awkwardly slow pace. After letting the researcher complete the drawn out question uninterrupted, the child then responded. For example, this exchange between the researcher and P14 reflected the common dynamic:

Researcher: Dooo yooouuu thhhiiinnnkkk weee shhhooouuullldddd trrreeaaattt huuummmaaannnsss annnddd rooobbbooottsss thhheee saaammmeee waaayyy?
P14: Yes.

Although children did not explicitly acknowledge the researcher's slowed speech, in some cases they showed nonverbal indicators of noticing the change. For example, in various instances children looked up sharply when the researcher began to speak slowly, began to smirk, laughed, grinned and made knowing eye-contact with their parent, or raised their eyebrows. These indicators suggest that the children may have noticed the slow speech, despite choosing not to engage with it. P19 illustrates this behavior in the following exchange:

Researcher: Dooo yooouuu evvveeerrr seeeeee yooouuurrr paaarrreeennntttsss—
P19: ((Stops fidgeting and stares intently at researcher))
Researcher: —orrr ottthhheeerrrsss usssiiinnnggg a diiiggiiitttaaalll asssssiiisssttaaannnttt?
P19: ((Pauses, looks at parent, and shrugs))

Similarly, P24 focused on and fidgeted with a plastic toy during the interview. He was holding it up to his eye when the researcher began speaking slowly, and on her fourth word, he set it down and looked up, staring at her. After a few more seconds of listening to her slow speech, a huge grin spread across his face. He glanced over at his mother, smiling, then looked back at the researcher and focused on her intently for the duration of her slow question, appearing to barely suppress laughter.

The one child (P11) who *did* confidently and immediately use the word *bungo* with the researcher had a particularly warm and playful interview session. She exuberantly chatted throughout the duration of the session, showing an outgoing side of her personality and no reluctance or hesitation to connect with this new person. Her responses were high-energy and fanciful, and she laughed loudly as she described wanting to own a CA that would make gifts for her, which she named "*Santa Claus 2*." She energetically pushed her chair around the room, laughed, and made jokes.

4.5 Sustained Transfer Beyond the Lab

According to parents' self-report, 11 of the 20 children whose parents tried the follow-up task at home spontaneously used the term *bungo* to ask their parent to speed up. Parents of these eleven children reported things like, "*When we got home i [sic] tried the same thing and [child name] actually used the word*" (Parent, P2) and:

> "*I DID remember to test this out later that evening and the next day…I would be talking with my daughter and then change to a slow pace and just like we did during our 'private' time in the study, she jumped in with a 'PONGO!' [sic] immediately as it happened*" (Parent, P17).

A number of parents commented on the immediacy and confidence of children's responses, saying the child interrupted or instantly used the trained keyword. For example, P15's parent described her child's reaction saying, "*[S]he immediately turned around with a smile and said 'bungo!' No hesitation at all.*" In contrast, nine of 20 children did not use the term *bungo* at home, and parents of these children said things like, "*[He] did not use the special word*" (Parent, P16).

Parents' descriptions of their children's responses at home were consistent with the themes we observed in the lab. Many parents described children responding to the term playfully and treating the interaction as an enjoyable shared experience. The parent of P23 and P24 said, "*They both used the word bongo [sic]…with a wry smile on their faces. Almost like it was a joke.*" Similarly, the parent of P1 reported that when her child responded with *bungo* he "*mentioned [it] with a smile, like it was our private joke (to my surprise!)*" The parent of P17 told us:

> "*I felt like it quickly became an inside joke. The next day I tested it with her during a three-way conversation with my husband. He didn't even notice my slowed pace of talking, but [child name] jumped in with 'Pongo' [sic] and then laughed that my husband doesn't understand us. Kind of insinuating that it was an inside joke.*"

Consistent with the results we saw in the lab, parents of three participants said their child expressed skepticism about their behavior. For example, one parent told us, "*[H]e just regarded me like I was crazy*" (Parent, P16), and another parent reported that her son told her to stop acting like a robot (Parent, P7). Much like the lab, parents of two participants reported that their child responded negatively, saying things like, "*[H]e was having none of it*" (Parent, P13).

5 DISCUSSION

5.1 Social Awareness in Children's Transfer

We found that participants quickly mastered a simple linguistic routine taught by a CA, abstracted the routine, and recognized its applicability in other contexts. Of greatest relevance to our research questions, we saw a strong suggestion that children differentiated how and when to apply the routine. When children interacted with CA2, their most common behavior was to transfer the routine and use it matter-of-factly. Although not all children spontaneously transferred *bungo* to this new context, most did, and all of them did so in a straightforward way. No child expressed skepticism about the appropriateness of using the routine in this context or refused to use it on principle. No child showed signs that they found it amusing or strange that CA2 responded to the linguistic routine taught by CA1.

When children encountered an opportunity to use the term *bungo* with a parent, their most common behavior was to do so, but unlike in their interaction with CA2, to do so playfully. Nearly as many children used the term in this context, but without the straightforward pragmatism that characterized their transfer to CA2. With no prompting to do so, a large fraction of parents used the phrase "*an inside joke*" to describe their child's reaction to using the keyword with them, either in the lab or at home.

Finally, when children encountered an opportunity to use the term *bungo* with a researcher they did not know, the most common response was to ignore the researcher's odd conversational behavior and to use pre-existing conversational norms that characterize person-to-person behavior. It is, of course, possible that children did not use the term *bungo* with the researcher because they did not recognize its applicability, and we do not have direct access to their interpretation of the conversation. However, just moments before, these same children used the term with their parents and explicitly commented on their parents' speech, suggesting they understand the pattern. The one child who quickly transferred *bungo* to her conversation with the researcher also showed an unusually high degree of warmth and outgoing exuberance throughout the session, and her playful use of *bungo* was consistent with other children's playful use of the term with their parents. Thus, we observed the following differentiation in children's behaviors:

- **Unfamiliar device (CA2) vs. familiar person (parent):** Children were equally likely to use the term *bungo* in these two contexts, but they used it as a necessary linguistic device when interacting with CA2 and as a shared joke when interacting with a parent.
- **Unfamiliar device (CA2) vs. unfamiliar person (researcher):** Both CA2 and the researcher were essentially strangers to the child, and we saw little evidence of children attempting to bond or create shared meaning in either context. Instead, children abided by the conversational norms of each context, using *bungo* with CA2 and not with the researcher.
- **Familiar person (parent) vs. unfamiliar person (researcher):** Children were far more likely to use *bungo* with a parent than with the researcher, and in each case,their behavior aligned with their pre-existing interpersonal relationship. Children playfully repurposed *bungo* as a joke when interacting with a parent but remained reserved and polite and did not use the routine when interacting with a stranger.

Children's willingness to experiment with a CA-trained routine with their parent and reluctance to do so with a stranger is consistent with prior work on attachment relationships. Children discriminate between their attachment figures (typically parents and other close caregivers) and strangers, and developmental literature shows children with secure attachments exhibit greater familiarity, intimacy, and risk-taking with parents than with strangers [18, 46, 59]. Children's use of *bungo* with parents may reflect a dual perspective that: 1) the term is only appropriate for CA interactions, but 2) its *inappropriate* use can be explored within the safe context of an attachment relationship.

5.2 An Opportunity for Building Valuable Linguistic Routines

Children's playfulness with parents suggests that CAs may be able to support children's development and wellbeing in exciting ways. The Joint Media Engagement framework (JME) explains that when shared digital experiences between parents and children contain key characteristics, they can become sources of meaning and growth for children [54]. For example, digital experiences that enable *boundary crossing* (by carrying over into daily life) and allow for *multiple planes of engagement* (such that children and parents can both participate actively) provide an excellent foundation for learning and connecting with parents. Here, we saw that children were willing to engage with parents around the *bungo* routine outside of the child-CA interaction context and to expand upon it playfully. This suggests CAs may be a useful tool for introducing constructive linguistic routines that parents can build on and reinforce.

For example, rather than teaching a nonsense term like *bungo*, a CA might teach a simple evidence-based linguistic routine for cultivating socioemotional growth. One existing curriculum teaches school-age children to take a "meta-moment" when they need to step outside an emotionally challenging situation and talk through it [10]. Like many socioemotional interventions, the meta-moment (and the larger curriculum that uses it) incorporates a number of linguistic routines that could be presented by a CA, such as positive self-talk, reframing, listing triggers, and articulating responses to structured prompts. All of these offer promise as simple routines that could be introduced and scaffolded by CAs and then expanded upon in interactions with parents.

Additionally, many other prior studies document linguistic routines that can help people cultivate strong relationships, such as paraphrasing a speaker's words with statements like, "*What I'm hearing you say is…*," explicitly labeling one's own feelings, waiting for a conversational opening to make an interjection (instead of interrupting), adding semantic softeners when making requests, and asking clarifying questions. A number of clinical interventions have been shown to be effective in helping children and parents cultivate these types of linguistic routines to enhance their social communication (e.g., [9, 23, 29, 42, 58]). Our results suggest these offline interventions have potential as CA interactions.

5.3 Separating People and CAs

Finally, prior work consistently shows that children readily distinguish CAs from people [66], and in keeping with this past literature, children in our study found it surprising when their parents behaved like CAs, sometimes reacting with skepticism or even resistance. Similarly, they behaved differently when they encountered slow speech from an unfamiliar CA and when they encountered slow speech from an unfamiliar person. This suggests that, despite children's interest in personified interfaces [25, 67], they simultaneously value the ontological distinction they draw between humans and machines. They may delight in social robots, CAs, and other human-like interfaces [44], but our participants signaled that these interfaces are distinct from people—particularly from the people they love. Their skepticism, playful reshaping, and resistance to blurring the line between person and agent suggests that it is important to them that this separation remain intact. This is consistent with prior work showing that children [68] and parents [17] alike express fears about personified technologies intruding on their attachment relationships.

Children's emphasis on separating parents and CAs suggests designers: 1) adopt a first-class design principle of preserving ontological differentiation between CA and human, and 2) build interfaces that proactively communicate this differentiation to the user. By implementing systems that present and reinforce the message that they are a class of entity distinct from people, these devices will more closely adhere to children's values and be more sensitive to their fears. This might involve answering social questions with responses that reflect the device's non-human status, adding *design seams* [15] that expose aspects of the CA's inner workings, adding surface-level detail to the interface to signal inanimacy (e.g., using a robotic voice rather than a human one), or numerous other design approaches.

5.4 Limitations and Future Work

Our investigation was conducted with a small group of children, in a one-time session, and with a single linguistic routine. The routine itself was highly contrived, and the effects we saw may differ with respect to patterns of speech that children also encounter in person-to-person interaction. The follow-up conversation at home was conducted without scripting or careful control. Children's behaviors might change with repeated exposure over time. Further, the patterns in children's behavior are likely to be influenced by changes to the design of the CA or the interaction. Both of our agents had light embodiment that many smart speaker CAs (like Alexa and Siri) lack, and prior work has shown that embodiment influences users' interaction behaviors (e.g., [37, 41]) and can increase engagement [38]. The awkwardness or novelty of the routine, the demeanor of the child, and whether the routine is explicitly trained or passively absorbed, might also play a role in children's behavior.

6 CONCLUSION

To examine whether CAs can influence children's conversational habits with people, we conducted an observational lab study in which children learned a new linguistic routine from a CA and then had the opportunity to transfer this routine to other contexts. Children both learned and transferred the new routine easily but their transfer behaviors differed by context. When conversing with another CA, children were most likely to transfer the routine matter-of-factly. When conversing with a parent, children were most likely to use the routine, but knowingly and with playfulness, bringing it into the conversation in a way that created an opportunity to bond and share an inside joke. When conversing with a stranger, they were most likely to set the routine aside and fall back on existing conversational norms. Our results suggest that CAs can change what children say and how they say it, but they also suggest children may be selective in when and how they apply these new habits. Children's quick mastery of a simple CA-trained routine and willingness to share it with parents suggests designers have an opportunity to work in partnership with families to create experiences that can serve as a source of shared meaning.

ACKNOWLEDGMENTS

Many thanks to all children and parents who participated in this study and to the anonymous reviewers for their feedback. We are indebted to Jon Froehlich, who provided excellent advice on an early draft of this manuscript. This work was supported in part by a Jacobs Foundation Early Career Fellowship to Alexis Hiniker.

SELECTION AND PARTICIPATION OF CHILDREN

Participants were recruited through social media, email lists, advertisements at schools, snowball sampling, and an institutional participant pool. Parents completed a consent process, where they were asked to participate as confederates. We asked parents to assess whether their participation as a confederate or the researcher's slowed speech might cause their child distress. We also told them that if they participated, they should discontinue at any time if they felt the procedure was causing their child distress. We began each session by securing the child's assent. We explained that we would like for them to talk to two CAs and that afterward we would ask them questions, and we asked if this would be ok. We submitted these procedures as a deception study to our IRB.

REFERENCES

[1] [n.d.]. The world's most advanced text to speech technology. https://www.cereproc.com/
[2] Mehdi Alaimi, Edith Law, Kevin Daniel Pantasdo, Pierre-Yves Oudeyer, and Hélène Sauzeon. 2020. Pedagogical Agents for Fostering Question-Asking Skills in Children. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
[3] Mohammad Rafayet Ali, Seyedeh Zahra Razavi, Raina Langevin, Abdullah Al Mamun, Benjamin Kane, Reza Rawassizadeh, Lenhart K Schubert, and Ehsan Hoque. 2020. A Virtual Conversational Agent for Teens with Autism Spectrum Disorder: Experimental Results and Design Lessons. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*. 1–8.
[4] Felix Ameka. 1987. A comparative analysis of linguistic routines in two languages: English and Ewe. *Journal of Pragmatics* 11, 3 (1987), 299–326.
[5] John Langshaw Austin. 1975. *How to do things with words*. Vol. 88. Oxford university press.
[6] Edward C. Baig. 2019. Say thank you and please: Should you be polite with Alexa and the Google Assistant? https://www.usatoday.com/story/tech/2019/10/10/do-ai-driven-voice-assistants-we-increasingly-rely-weather-news-homework-help-otherwise-keep-us-info/3928733002/
[7] Judith A Becker. 1981. Preschoolers' Judgments of Speaker Status Based on Requests. (1981).
[8] Judith A Becker. 1986. Bossy and nice requests: Children's production and interpretation. *Merrill-Palmer Quarterly (1982-)* (1986), 393–413.
[9] Douglas Biklen, Mary W Morton, Deborah Gold, Carol Berrigan, and Sudha

Swaminathan. 1992. Facilitated communication: implications for individuals with autism. *Topics in language disorders* (1992).

[10] Marc A Brackett, Susan E Rivers, Maria R Reyes, and Peter Salovey. 2012. Enhancing academic performance and social and emotional competence with the RULER feeling words curriculum. *Learning and Individual Differences* 22, 2 (2012), 218–224.

[11] Cynthia Breazeal, Paul L. Harris, David DeSteno, Jacqueline M. Kory Westlund, Leah Dickens, and Sooyeon Jeong. 2016. Young Children Treat Robots as Informants. *Topics in Cognitive Science* 8, 2 (2016), 481–491.

[12] Nathan G Burton and James Gaskin. 2019. "Thank You, Siri": Politeness and Intelligent Digital Assistants. (2019).

[13] Katja F Cantone. 2007. *Code-switching in bilingual children.* Vol. 296. Springer.

[14] Fabio Catania, Micol Spitale, Giulia Cosentino, and Franca Garzotto. 2020. Conversational Agents to Promote Children's Verbal Communication Skills. In *International Workshop on Chatbot Research and Design.* Springer, 158–172.

[15] Matthew Chalmers and Ian MacColl. 2003. Seamful and seamless design in ubiquitous computing. In *Workshop at the crossroads: The interaction of HCI and systems issues in UbiComp*, Vol. 8.

[16] Catherine C Chase, Doris B Chin, Marily A Oppezzo, and Daniel L Schwartz. 2009. Teachable agents and the protégé effect: Increasing the effort towards learning. *Journal of Science Education and Technology* 18, 4 (2009), 334–352.

[17] Ying-Yu Chen, Ziyue Li, Daniela Rosner, and Alexis Hiniker. 2019. Understanding Parents' Perspectives on Mealtime Technology. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 1 (2019), 1–19.

[18] Kim Chisholm, Margaret C Carter, Elinor W Ames, and Sara J Morison. 1995. Attachment security and indiscriminately friendly behavior in children adopted from Romanian orphanages. *Development and psychopathology* 7, 2 (1995), 283–294.

[19] Eve V Clark. 2009. *First language acquisition.* Cambridge University Press.

[20] Florian Coulmas. 2011. *Conversational routine: Explorations in standardized communication situations and prepatterned speech.* Vol. 96. Walter de Gruyter.

[21] Mike Elgan. 2018. The case against teaching kids to be polite to Alexa. https://www.fastcompany.com/40588020/the-case-against-teaching-kids-to-be-polite-to-alexa

[22] Goffman Erving. 1971. Relations in public: Microstudies of the public order.

[23] Sheila M Eyberg, Stephen R Boggs, and James Algina. 1995. Parent-child interaction therapy: a psychosocial model for the treatment of young children with conduct problem behavior and their families. *Psychopharmacology bulletin* (1995).

[24] Radhika Garg and Subhasree Sengupta. 2019. " When you can do it, why can't I?": Racial and Socioeconomic Differences in Family Technology Use and Non-Use. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–22.

[25] Radhika Garg and Subhasree Sengupta. 2020. Conversational Technologies for In-home Learning: Using Co-Design to Understand Children's and Parents' Perspectives. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* 1–13.

[26] Radhika Garg and Subhasree Sengupta. 2020. He Is Just Like Me: A Study of the Long-Term Use of Smart Speakers by Parents and Children. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 1 (2020), 1–24.

[27] Jean Berko Gleason, Rivka Y Perlmann, and Esther Blank Greif. 1984. What's the magic word: Learning language through politeness routines. *Discourse processes* 7, 4 (1984), 493–502.

[28] G. Gordon, C. Breazeal, and S. Engel. 2015. Can Children Catch Curiosity from a Social Robot?. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI).* 91–98.

[29] Steven E Gutstein, Audrey F Burgess, and Ken Montfort. 2007. Evaluation of the relationship development intervention program. *Autism* 11, 5 (2007), 397–411.

[30] Christian Heath, Jon Hindmarsh, and Paul Luff. 2010. *Video in qualitative research.* Sage Publications.

[31] Juan Pablo Hourcade, Natasha E Bullock-Rest, and Thomas E Hansen. 2012. Multitouch tablet applications and activities to enhance the social skills of children with autism spectrum disorders. *Personal and ubiquitous computing* 16, 2 (2012), 157–168.

[32] Gail Jefferson. 2004. Glossary of transcript symbols with an introduction. *Pragmatics and Beyond New Series* 125 (2004), 13–34.

[33] Peter H. Kahn Jr, Heather E. Gary, and Solace Shen. 2013. Children's Social Relationships With Current and Near-Future Robots. *Child Development Perspectives* 7, 1 (2013), 32–37. https://srcd.onlinelibrary.wiley.com/doi/abs/10.1111/cdep.12011

[34] Takayuki Kanda, Takayuki Hirano, Daniel Eaton, and Hiroshi Ishiguro. 2004. Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. *Human–Computer Interaction* 19, 1-2 (2004), 61–84.

[35] Elizabeth S. Kim, Lauren D. Berkovits, Emily P. Bernier, Dan Leyzberg, Frederick Shic, Rhea Paul, and Brian Scassellati. 2013. Social Robots as Embedded Reinforcers of Social Behavior in Children with Autism. *Journal of Autism and Developmental Disorders* 43, 5 (01 May 2013), 1038–1049. https://doi.org/10.1007/s10803-012-1645-2

[36] Zoe Kleinman. 2018. Alexa, are you friends with our kids? *BBC News* (July 2018). https://www.bbc.com/news/technology-44847184

[37] Dimosthenis Kontogiorgos, Gabriel Skantze, André Tiago Abelho Pereira, and Joakim Gustafson. 2019. The effects of embodiment and social eye-gaze in conversational agents. In *41st Annual Meeting of the Cognitive Science (CogSci), Montreal July 24th–Saturday July 27th, 2019.*

[38] Hatice Kose-Bagci, Ester Ferrari, Kerstin Dautenhahn, Dag Sverre Syrdal, and Chrystopher L Nehaniv. 2009. Effects of embodiment and gestures on social interaction in drumming games with a humanoid robot. *Advanced Robotics* 23, 14 (2009), 1951–1996.

[39] Deborah L Linebarger and Dale Walker. 2005. Infants' and toddlers' television viewing and language outcomes. *American behavioral scientist* 48, 5 (2005), 624–645.

[40] 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.

[41] Michal Luria, Samantha Reig, Xiang Zhi Tan, Aaron Steinfeld, Jodi Forlizzi, and John Zimmerman. 2019. Re-Embodiment and Co-Embodiment: Exploration of social presence for robots and conversational agents. In *Proceedings of the 2019 on Designing Interactive Systems Conference.* 633–644.

[42] Johnny L Matson, Mary E Shoemaker, Megan Sipes, Max Horovitz, Julie A Worley, and Alison M Kozlowski. 2011. Replacement behaviors for identified functions of challenging behaviors. *Research in Developmental Disabilities* 32, 2 (2011), 681–684.

[43] Amy Ogan, Samantha Finkelstein, Erin Walker, Ryan Carlson, and Justine Cassell. 2012. Rudeness and rapport: Insults and learning gains in peer tutoring. In *International Conference on Intelligent Tutoring Systems.* Springer, 11–21.

[44] Hae Won Park, Rinat Rosenberg-Kima, Maor Rosenberg, Goren Gordon, and Cynthia Breazeal. 2017. Growing Growth Mindset with a Social Robot Peer. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (Vienna, Austria) *(HRI '17).* Association for Computing Machinery, New York, NY, USA, 137–145. https://doi.org/10.1145/2909824.3020213

[45] Leo Natan Paschoal, Lucas Fernandes Turci, Tayana Uchôa Conte, and Simone RS Souza. 2019. Towards a Conversational Agent to Support the Software Testing Education. In *Proceedings of the XXXIII Brazilian Symposium on Software Engineering.* 57–66.

[46] Katherine C Pears, Jacqueline Bruce, Philip A Fisher, and Hyoun K Kim. 2010. Indiscriminate friendliness in maltreated foster children. *Child maltreatment* 15, 1 (2010), 64–75.

[47] Martin Porcheron, Joel E Fischer, Stuart Reeves, and Sarah Sharples. 2018. Voice interfaces in everyday life. In *proceedings of the 2018 CHI conference on human factors in computing systems.* 1–12.

[48] Simon Provoost, Ho Ming Lau, Jeroen Ruwaard, and Heleen Riper. 2017. Embodied conversational agents in clinical psychology: a scoping review. *Journal of medical Internet research* 19, 5 (2017), e151.

[49] Mabel Rice. 1983. The role of television in language acquisition. *Developmental Review* 3, 2 (1983), 211–224.

[50] Jessica Schroeder, Chelsey Wilkes, Kael Rowan, Arturo Toledo, Ann Paradiso, Mary Czerwinski, Gloria Mark, and Marsha M Linehan. 2018. Pocket skills: A conversational mobile web app to support dialectical behavioral therapy. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* 1–15.

[51] Rachel L Severson and Stephanie M Carlson. 2010. Behaving as or behaving as if? Children's conceptions of personified robots and the emergence of a new ontological category. *Neural Networks* 23, 8-9 (2010), 1099–1103.

[52] Catherine E Snow and Beverly A Goldfield. 1983. Turn the page please: Situation-specific language acquisition. *Journal of Child language* 10, 3 (1983), 551–569.

[53] Micol Spitale. 2019. Training self-sufficiency and social skills with embodied conversational agent for children with autism. In *Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion.* 151–152.

[54] Reed Stevens and L Takeuchi. 2011. The new coviewing: Designing for learning through joint media engagement. (2011).

[55] Fumihide Tanaka and Shizuko Matsuzoe. 2012. Children Teach a Care-receiving Robot to Promote Their Learning: Field Experiments in a Classroom for Vocabulary Learning. *J. Hum.-Robot Interact.* 1, 1 (July 2012), 78–95.

[56] Hiroki Tanaka, Hideki Negoro, Hidemi Iwasaka, and Satoshi Nakamura. 2017. Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders. *PloS one* 12, 8 (2017).

[57] Andrea Tartaro and Justine Cassell. 2008. Playing with virtual peers: bootstrapping contingent discourse in children with autism.. In *ICLS (2).* Citeseer, 382–389.

[58] Rae Thomas and Melanie J Zimmer-Gembeck. 2007. Behavioral outcomes of parent-child interaction therapy and Triple P—Positive Parenting Program: A review and meta-analysis. *Journal of abnormal child psychology* 35, 3 (2007), 475–495.

[59] Barbara Tizard. 1977. *Adoption: A second chance.* Free Press.

[60] Michael Tomasello, Michael Jeffrey Farrar, and Jennifer Dines. 1984. Children's speech revisions for a familiar and an unfamiliar adult. *Journal of Speech, Language, and Hearing Research* 27, 3 (1984), 359–363.

[61] Alice Truong. 2016. Parents are worried the Amazon Echo is conditioning their kids to be rude. https://qz.com/701521/parents-are-worried-the-amazon-echo-is-conditioning-their-kids-to-be-rude/

[62] Michael Villano, Charles R. Crowell, Kristin Wier, Karen Tang, Brynn Thomas, Nicole Shea, Lauren M. Schmitt, and Joshua J. Diehl. 2011. DOMER: A Wizard of Oz Interface for Using Interactive Robots to Scaffold Social Skills for Children with Autism Spectrum Disorders. In *Proceedings of the 6th International Conference on Human-robot Interaction* (Lausanne, Switzerland) *(HRI '11)*. ACM, New York, NY, USA, 279–280. https://doi.org/10.1145/1957656.1957770

[63] Iain Werry, Kerstin Dautenhahn, Bernard Ogden, and William Harwin. 2001. Can Social Interaction Skills Be Taught by a Social Agent? The Role of a Robotic Mediator in Autism Therapy. In *Cognitive Technology: Instruments of Mind*, Meurig Beynon, Chrystopher L. Nehaniv, and Kerstin Dautenhahn (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 57–74.

[64] Rainer Winkler, Sebastian Hobert, Antti Salovaara, Matthias Söllner, and Jan Marco Leimeister. 2020. Sara, the lecturer: Improving learning in online education with a scaffolding-based conversational agent. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[65] Beverly Park Woolf, Ivon Arroyo, Kasia Muldner, Winslow Burleson, David G Cooper, Robert Dolan, and Robert M Christopherson. 2010. The effect of motivational learning companions on low achieving students and students with disabilities. In *International Conference on Intelligent Tutoring Systems*. Springer, 327–337.

[66] 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*. 1–13.

[67] Svetlana Yarosh, Stryker Thompson, Kathleen Watson, Alice Chase, Ashwin Senthilkumar, Ye Yuan, and AJ Bernheim Brush. 2018. Children asking questions: speech interface reformulations and personification preferences. In *Proceedings of the 17th ACM Conference on Interaction Design and Children*. 300–312.

[68] Jason C Yip, Kiley Sobel, Xin Gao, Allison Marie Hishikawa, Alexis Lim, Laura Meng, Romaine Flor Ofiana, Justin Park, and Alexis Hiniker. 2019. Laughing is Scary, but Farting is Cute: A Conceptual Model of Children's Perspectives of Creepy Technologies. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–15.