

Communication Breakdowns Between Families and Alexa

Erin Beneteau

Information School
University of Washington
Seattle, WA, USA
ebenet@uw.edu

Olivia K. Richards

Pennsylvania State University
University Park, PA, USA
orichards@psu.edu

Mingrui Zhang

Information School
University of Washington
Seattle, WA, USA
mingrui@uw.edu

Julie A. Kientz

Human Centered Design and
Engineering
University of Washington
Seattle, WA, USA
jkientz@uw.edu

Jason Yip

Information School
University of Washington
Seattle, WA, USA
jcyip@uw.edu

Alexis Hiniker

Information School
University of Washington
Seattle, WA, USA
alexisr@uw.edu

ABSTRACT

We investigate how families repair communication breakdowns with digital home assistants. We recruited 10 diverse families to use an Amazon Echo Dot in their homes for four weeks. All families had at least one child between four and 17 years old. Each family participated in pre- and post- deployment interviews. Their interactions with the Echo Dot (Alexa) were audio recorded throughout the study. We analyzed 59 communication breakdown interactions between family members and Alexa, framing our analysis with concepts from HCI and speech-language pathology. Our findings indicate that family members collaborate using discourse scaffolding (supportive communication guidance) and a variety of speech and language modifications in their attempts to repair communication breakdowns with Alexa. Alexa's responses also influence the repair strategies that families use. Designers can relieve the communication repair burden that primarily rests with families by increasing digital home assistants' abilities to collaborate together with users to repair communication breakdowns.

CCS CONCEPTS

• Human-centered computing → **Human computer interaction (HCI)**; *Collaborative interaction, Field studies, Empirical studies in HCI*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
CHI 2019, May 4–9, 2019, Glasgow, Scotland UK
© 2019 Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-5970-2/19/05...\$15.00
<https://doi.org/10.1145/3290605.3300473>

KEYWORDS

Digital Home Assistants; Human Computer Interaction; Joint Media Engagement; Communication Repairs; Families

ACM Reference format:

Erin Beneteau, Olivia K. Richards, Mingrui Zhang, Julie A. Kientz, Jason Yip, and Alexis Hiniker. 2019. Communication Breakdowns Between Families and Alexa. In *2019 CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland, UK*. ACM, New York, NY, USA. 13 pages.
<https://doi.org/10.1145/3290605.3300473>

1 INTRODUCTION

The worlds of human-to-human communication and human-to-technology communication are quickly becoming blurred in everyday, family life. The increasing popularity of digital home assistants, like the Amazon Echo, and conversational assistants, like Siri, increase users' expectations of voice as an effective communication method with machines [23]. However, humans must work to adapt their communication patterns to the needs of the machines, rather than machines adapting to humans [19]. People shorten their sentences, use simplified language, and repeat themselves in attempts to be understood by voice interfaces [19,20,22,23,31]. As a result, users of voice interface technology become frustrated and can fail to learn the full capabilities of the technology, or abandon use altogether [10,23].

Despite the difficulty with human-to-voice interface communication, the technology is becoming pervasive in family lives. Yarosh et. al found that 93% of the children they talked with at a state fair in Minnesota had used voice interface technology [43]. Digital home assistants, which utilize voice interfaces, are anticipated to be present in 55% of American households by the year 2022 [32]. Yet,

the prevalence of the technology does not necessarily indicate the successful, ongoing use of the technology. True conversational capabilities have not yet been fully realized [33,34], and we want to better understand how human-technology “conversations” can be improved.

In our research, we blend two fields of study: human-computer interaction (HCI) and speech-language pathology. HCI has a long history of research in how humans communicate with computers, robots, and other conversational agents [26,39]. Speech language pathology has a long history of research on human-to-human communication development, communication disorders, and remediation [14,35]. Drawing from these two fields, we aim to better understand the communication breakdowns and repair strategies used between families and digital home assistants, specifically, the Amazon Echo Dot (who we refer to as the conversational partner “Alexa”).

To frame our study, we use the concepts of joint media engagement [40], conversational analysis [27], pragmatics [35], and discourse scaffolding [41]. Building on these concepts, we explore the communicative relationships between digital home assistants and families. Our specific research questions are: 1) What types of communication breakdowns occur? 2) How do families repair communication breakdowns with Alexa? 3) How does Alexa repair communication breakdowns with family members?

We gave Amazon Echo Dots to 10 economically and ethnically diverse families to use in their homes for four weeks. None of the families had previously owned a digital home assistant. We conducted pre- and post-deployment interviews with all 10 families in their homes, and we captured audio of natural, unscripted family interactions with Alexa. Based on an analysis of 59 conversational interactions, we discuss three types of responses from Alexa that signal communication breakdowns and a taxonomy of five strategies that family members used to repair breakdowns. At times, Alexa initiated collaborative communication repair opportunities with her communication partner, but currently, the brunt of communication repair work lies with the family. Designers could use the analysis of communication repair types to create relevant and helpful responses for digital home assistants to use during communication breakdowns. We conclude with design suggestions that can improve digital home assistants’ facilitation of conversational repairs with families.

2 RELATED WORK

Human Communication and Communication Breakdowns

In the field of speech-language pathology, pragmatics incorporates the social aspects of communication, including turn taking, topic maintenance, socially appropriate speech and language use, and code switching [18,35]. Code switching involves the recognition of different communicative expectations for different communication partners and/or different communication environments [8]. The most explicit form of code switching is to switch languages to match the language of the communication partner, such as shifting from English to Spanish [13]. A subtle form of code-switching, in which the rules and use of language are dependent on the communication context, is demonstrated when a teenager uses a different conversational tone and different vocabulary when talking to friends than when talking to a parent. Pragmatics, and specifically code switching, requires some understanding and intentional thought behind communication interactions [17,35,36]. As such, linguistic code switching often begins around ages 4-5, when children begin intentionally talking in different ways to different people, in different settings [44].

Communication repair refers to the work of restoring shared understanding after conversational partners misunderstand each other. Essentially, the person who is talking needs to rephrase or say something differently because the person they were talking to did not understand what they were saying [25]. The ability to code switch can aid in conversational repair and can be thought of as the act of adjusting your speaking style to accommodate the listener [41]. These adjustments to speaking style can take many forms and are dependent on the communicator’s language abilities. For example, children often use repetition as their initial repair strategy when presented with “neutral” clarification responses, such as “huh,” “what,” and “I don’t know” [6,25]. Children ages nine and older might use contextual cues, such as defining terms, to repair a communication breakdown, whereas this can be more challenging for younger children [6]. Adults may engage in discourse scaffolding to assist children in expanding or adjusting their speech [41]. Discourse scaffolding is a mechanism to guide learning through communication strategies with the intention of gradually transferring skills and responsibility to the learner [38,41]. At an early communication stage, an example of discourse scaffolding is when a child says,

“cookie,” and a parent responds with an expansion of the child’s utterance: “*you want more cookies.*”

Dynamics between multiple communication partners, especially those involving young children, can be heavily influenced by the communicative development ages of the communication partners and the context of the communication interaction [14,18,21,24]. Our earlier example of discourse scaffolding was appropriate for a very young child, whereas the communication interaction between a teenager and parent would look very different. For example, a parent might use discourse scaffolding through questioning to lead the teen to expand on their own language. The parent might ask, “*What did you do after school?*” and the teen might reply, “*hung out with friends.*” The parent could then scaffold further elaboration with questions, such as, “*where did you and your friends hang out?*” While these examples highlight the nuances of communication interactions between family members of different ages and developmental stages, we find that the communication interaction can become even more complex when one of the communication partners is not human.

Human-Computer Communication

The field of HCI has a well-established body of literature on how humans verbally communicate with computers, robots, and other devices. In 1987, Suchman framed acts of human-machine interaction as a dialog between communication partners [39]. From this perspective, the work of the designer is to enable human and machine collaboration towards a shared understanding through continuous acts of collaborative communication repair when breakdowns occur. Yet, despite this legacy, and improvements in human-computer communication techniques [2–4], research continues to demonstrate that humans adapt their communication styles and patterns to match the machine, both with robots [31,42] and computers [28,30], rather than the other way around. Humans shorten their sentences [20,31], use repetition [1], increase volume [5] and hyperarticulate [28] as repair strategies. These modification strategies are motivated by a desire to achieve successful communication with computers [5,30].

Communication Breakdowns with Voice Interfaces

In recent years, the field of HCI has begun to examine human communication with conversational agents, such as Siri, and with digital home assistants, such as the Amazon Echo [10,23,34,37]. Despite the “conversational” interface with conversational agents, people are not yet

able to talk to technology in the same way that they talk to other people [23,34]. Users of conversational assistants often need to shorten their queries to key words, since increased utterance length can increase the likelihood of speech recognition errors, both with conversational agents and with other humans [19,22–24]. We also see young children using repetition as an initial conversational repair strategy when they are talking to a computer game [9]. With current systems, the burden of ensuring a successful communication interaction with a conversational agent continues to fall to the human in the conversation, with little support from the conversational agent itself [12,34].

The majority of research on communication breakdowns and repairs takes place in 1:1 interactions [19,20,22,24,25,31,43]. One of the few studies that investigates how multiple people work together to repair communication breakdowns with a conversational agent observed friends meeting in a café, using the conversational agents (such as Siri) on their phones [34]. The researchers used conversational analysis as their approach, based on the principle of analyzing naturally occurring interactions with technology from the user’s point of view [27]. Their findings revealed that when a communication breakdown occurs with a conversational agent, multiple people attempt communication repairs, passing the phone from person to person. Our specific interest is how families repair communication breakdowns with digital home assistants in naturalistic use in their homes.

We can turn to literature on how technology is situated within family dynamics for reference, such as Takeuchi and Stevens documentation of “Joint Media Engagement” (JME) and the ways families make meaning around their technology experience [40]. Family members, particularly parents, use media as shared learning opportunities, scaffolding communication around the media experience [16,40]. Comparatively little work has examined children’s and families’ interactions with digital assistants. Yarosh and colleagues found that children struggled to formulate queries to a voice interface [43] and Cheng and colleagues define a taxonomy of communication repair strategies that preschoolers use when talking to a voice-driven game [9]. Porcheron et al. report on a study of five families using the Amazon Echo in their homes, and they describe how a family of two adults and two children take turns in attempting to repair a communication breakdown with their Amazon Echo [33]. The authors conclude that the device’s responses to the communication breakdown are

ineffective in supporting the participants in identifying the cause of the breakdown.

In this study, we expand the body of work on communication repair strategies with conversational agents and digital home assistants. We specifically explore how diverse families attempt to repair communication breakdowns with the conversational partner Alexa (their Echo Dot) through a close analysis of conversational interactions.

3 METHOD

We conducted a four-week field study with 10 families, audio recording their everyday interactions with the Amazon Echo Dot (second generation).

Participants

We recruited 10 diverse families to use the Echo Dot for a period of four weeks, intentionally recruiting families who represent a wide-spectrum of family life. Families self-identified as never having owned a digital home assistant. Families also reported having a total household income at or below the median for the county in which the study was conducted. Families ranged from families of two to families of five (see Table 1). All families had at least one child between the ages of four and 17 living in the home. Two families were bilingual, in which all family members spoke both English and Spanish in their homes.

Procedure

We collected audio data throughout the four-week period, using a custom-built audio sampling system designed to record interactions every time the trigger word “Alexa” was spoken. We also conducted pre- and post-deployment interviews with families in their homes at the start of the deployment period and at the conclusion of the four weeks. We compensated families with US\$100 in gift cards for completing the study, and they were able to keep their Echo Dot after the conclusion of the study. The study was approved by our university’s Institutional Review Board.

Table 1. Total family members, ages, and ethnicity of participant families.

Family ID	N	Adult N	Adults age ranges	Child ages	Ethnicity
A	4	3	41-68+	13	Asian
B	4	3	18-55	12	White/Asian
C	2	1	41-55	9	White
D	4	2	26-55	3, 5	Hispanic
E	5	3	26-67	4, 8	White

Family ID	N	Adult N	Adults age ranges	Child ages	Ethnicity
F	5	3	18-55	<1, 5	Hispanic
G	3	1	26-40	10, 16	African American/ White
H	2	1	41-55	9	White
I	5	2	26-40	2, 4, 6	White
J	2	1	18-25	4	White

Audio Capture

To capture audio recordings of participants’ interactions with the Echo Dot, we created an audio buffer system inspired by the open source code from Porcheron et al. [33]. The software captures audio when the trigger word “Alexa” is spoken; it saves one minute of audio prior to the trigger word, along with three additional minutes during and after the trigger word is spoken. It then pushes a total of four minutes of audio to a secure server.

For our study, we deployed the audio recording system on a Samsung tablet computer that families placed near the Echo Dot in each of their homes. We instructed families to keep the tablet plugged in and not to use the tablet for anything other than for the study’s audio recording. The recording system also had both a “record a thought” button and a “delete” button available for families to leave explicit comments for the researchers or to delete the audio capture files, based on the desires of the participants.

Analysis

We used an inductive process to analyze the audio capture conversational samples [7,11,29]. We began with memoing and open coding during the initial transcriptions of the audio capture files. Through memoing and open coding, we noticed an emerging theme related to communication breakdowns. We then began coding communication breakdowns from transcripts of each of the 10 families, developing and revising codes as we found additional examples of communication breakdowns, reviewing a total of 14.5 hours of audio capture. We continued this process until communication breakdown codes were stable and applicable to multiple families. Once codes were stable, we reviewed transcripts from each of the 10 families for communication breakdowns again. We included communication breakdowns from each family in our corpus of 59 communication breakdown interactions, systematically going through each individual family’s transcript and pulling out communication breakdowns for each code (when present). If a family had multiple

communication breakdowns for a particular code type, we reviewed the instances and chose no more than three instances of that code type for deeper analysis. For example, Family C had multiple instances of adjusting prosody; we chose three instances that we felt were most representative to be included in the corpus for deeper analysis.

For our final analysis on communication breakdowns and repairs, a total of 59 conversational interactions falling under the broad themes of communication breakdowns and repairs were deeply analyzed by two researchers. Communication breakdowns were defined as interactions between family members and Alexa which did not result in an appropriate response from Alexa to the human communication partner. We drew on the HCI conversational analysis approach to analyze family conversations set in natural environments, with a focus on the user’s experiences. The use of conversational analysis in HCI aims to highlight opportunities for designers to improve interactions between humans and computers by detailing the structure of interactions from the user’s perspective [27].

The first author, a speech-language pathologist (SLP), used speech-language constructs when analyzing communication breakdowns and repairs. Constructs are defined in our findings below. We also consulted a practicing SLP on the codes and themes related to family communication repair strategies. The practicing SLP validated the communication repair strategies identified and validated the repair types used in the conversational samples discussed in this paper. For additional validation, we consulted a second SLP, also a researcher, who validated the repair strategy definitions used.

During our analysis, we discovered that family members were not the only communication partners to attempt communication repairs in conversations. Alexa also engaged in communication repair attempts, and therefore we have included an analysis of Alexa’s communication repairs in the findings, in addition to family communication repairs.

We also want to explicitly state that in our findings and discussion, we refer to Alexa in a personified form. We do this for two reasons: 1) Alexa is a communication partner in the conversational interactions and is most easily identified by name and personal pronoun when discussing the complex interaction between multiple conversation partners. 2) We reflect our participant families’ language describing Alexa, based on our interviews and audio capture.

4 RESULTS

All 10 of our families experienced communication breakdowns with Alexa. To illustrate our findings, we discuss three distinct conversational samples from three different families in which we highlight the complex communication dynamics between multiple family members and Alexa. Before our discussion of these rich conversational samples, we provide an explanation of the constructs we use to describe the communication interactions between family members and Alexa (shown in Tables 2-4).

Our families used a variety of specific speech and language repair strategies, which we categorize under the broader construct of “communication repair types.” Communication repair types are listed in Table 2, with quotes from our participant families used as examples.

In addition to specific repair strategies, we define a variety of discourse scaffolds between family members that occur during collaboration, adapted from the fields of education and speech-language pathology [38,41]. These scaffolds occur both between family members and between family members and Alexa. The discourse scaffolds are listed and defined in Table 3. Examples of discourse scaffolds are described when we discuss the three specific family communication breakdowns.

Table 2. Speech and language repair strategies used by family members.

Repair Type	Definition	Example
Prosodic changes	Adjustments to the rhythm or cadence of speech, including pausing and the rate of speech	“Alexa . . . what is . . . the . . . temperature?” [each individual word pronounced slowly and clearly instead of in a conversational manner]- Family C
Overarticulation	Exaggerating sounds, also referred to as hyperarticulation [22]	“Alexa, play ‘Make it Rain’ by Dack steN” [emphasis and prolongation of final consonant]- Family C
Semantic adjustments and modification	Modifying the meaning of a word or sentence, including providing “cues,” such as defining a word [6]	“Alexa, play little kid music.” (Alexa): I couldn’t find any little kid songs. “Alexa, play kids songs.”- Family G
Increased volume	The speaker raises their voice specifically for the interaction with Alexa	“Alexa stop.” “Alexa stop!” [louder] - Family D

Repair Type	Definition	Example
Syntactical adjustments	Modifying sentence structure or grammatical elements	“Alexa, thank you, stop it.” (alarm continues) “Alexa, stop.”- Family H
Repetition	The speaker repeats an utterance used earlier in the same conversation	“Alexa, call [name]. (Alexa): Do you mean [name]? “Volume to eight. Call [name]. Alexa, call [name]”.- Family A

Our final construct is that of Alexa’s responses and how they signal communication breakdowns. We encountered three types of responses from Alexa that indicate a communication breakdown. Examples from our audio capture of Alexa’s responses are listed in Table 4.

Table 3. Discourse scaffolds families used to support one another in reformulating their communication with Alexa.

Discourse Scaffold	Definition
Direct instruction	Telling a family member what they should say or why something has happened
Modeling	Producing an utterance to demonstrate the desired response
Redirection	Refocusing the conversation on a desired topic
Expansion	Adding on to something said by someone else
Contraction	Shortening something that has been said
Consulting	Family members asking others for assistance or information

Table 4. Signals from Alexa that a communication breakdown has occurred.

Response type	Definition	Example
Acting on Misunderstanding (AoM)	Performing an action or providing a response based on misheard or misunderstood input	“Alexa, what should we do this night?” Alexa responds with a definition of “this night.”
Neutral Clarification Response (NR)	Providing an indication that the communication partner’s interaction was unclear	“Sorry, I don’t know that.” “Sorry, I’m not sure.” “I’m not quite sure how to help you with that.”

Response type	Definition	Example
Specific Clarification Response (SR)	Providing specific information to request clarification	“Did you mean [person’s name]?” “Is that 10:45 in the morning or evening?”

Family Collaboration on Repair Strategies

We explored the variety of communication repair strategies that family members use with Alexa, within the framing of family collaboration and joint media engagement. We present three examples of communication breakdowns with three different families, and use speech and language constructs to analyze the specific repair strategies used in detail.

Communication Breakdown 1: Family E

Our first communication breakdown occurs in Family E, who are talking while music is playing on their Echo Dot. The topic of conversation is about popsicles, a flavored ice treat. The son (age 8) initiates communication with Alexa. Alexa responds by acting on a misunderstanding. The mother collaborates with her son in repairing the communication breakdown, to which Alexa responds with specific clarification responses. The family ultimately carries on the conversational topic without Alexa.

(Son): Can I ask a question?
 (Mother): What?
 (Father): What’s the question?
 (Son): Um (pause) what are popsicles made of?
 (Father): Why would I tell you that?
 (Son): Alexa, (pause and music stops) what are popsicles made of?
 (Alexa): Three, six servings of s’mores popsicles are, you’ll need two cups cold milk, one . . . (Alexa continues to talk)
 (Mother, talking over Alexa): She thinks it’s s’mores popsicles because you didn’t articulate.
 (Mother, in louder voice): Alexa, (pause and Alexa stops talking) how do you make a popsicle?
 (Alexa): Okay, for popsicles, I recommend a top recipe called sweet pink popsicle orange summers, which take eight hours and five minutes to make.
 (Mother talking over Alexa): Oh my god.
 (Alexa): You can ask for more information, or for more recipes, say next.
 (Son): Next
 (Mother): Alexa, stop. Resume music.
 (Father): Hey. You’re a smart kid. What do you think popsicles are made of?
 (Son): Ice?
 (Father): No.
 (Mother): How do you make ice?
 (Father): How do you make it?

(Son): Water.
 (Mother): Plus what?
 (Son): Plus steam! No, plus cold.
 (conversation continues between parents and child about
 popsicles)

In this interaction, the son first asks his parents how to make popsicles, asking for consultation without the use of media. When he does not get the answer he is looking for, he changes communication partners and asks Alexa. At this point, the family as a unit is now engaged with media, with Alexa as a consultant. We see throughout the interaction that all family members are listening to the conversational interactions with Alexa, as evidenced by the subsequent communication by family members related to Alexa's responses.

We know the son wants generalized information based on his initial word choice, "*what are popsicles made of?*" The use of a plural and the lack of a specific type of popsicle indicates that he is interested in popsicles in general instead of a specific popsicle type.

Alexa responds on the correct topic—popsicles—but provides too specific of a response that deviates from the son's intended topic. When the communication breakdown occurs, the son's mother actively engages in the interaction. However, her initial communication is with her son rather than with Alexa. She uses the communication breakdown as a learning opportunity to educate her son on improving his success in obtaining information from Alexa. The mother provides discourse scaffolding in the form of direct instruction. She instructs her son, given her interpretation of the source of the communication breakdown with Alexa: his poor articulation in which Alexa misheard the word "*s'mores.*" She then provides further discourse scaffolding in the form of modeling by engaging with Alexa directly in a communication repair.

The mother makes the first repair attempt by adjusting the syntax of the question from, "*what are popsicles made of?*" to, "*how do you make a popsicle.*" As a result, the communication breakdown is somewhat repaired, and Alexa responds both on-topic and with specific information relevant to the question. However, Alexa continues to provide information for a specific type of popsicle rather than a generalized recipe for any popsicle. At this point, Alexa assists her communication partners by refining her response and providing a summary of the popsicle recipe. Alexa then models two possible communication responses to her communication partner: "*ask for more information*" or "*say next.*" By providing choices, Alexa is performing discourse scaffolding of her

own and promoting turn-taking in the communication interaction through a specific clarification response. The specific clarification further refines her response to more closely match the needs of her communication partner.

The son, the original instigator of the communication interaction, resumes his role in the interaction with Alexa and responds to Alexa with "*next,*" one of the communicative options that Alexa provided. However, it is clear that both the mother and father are engaged with the interaction and have had enough of Alexa's responses. The mother redirects the conversation by stopping Alexa. The parents take on the roles of communication partners with their son. The family maintains the topic of conversation but abandons Alexa as a communication partner for the remainder of the conversational topic. The parents resume the discourse scaffolding that the mother started earlier and lead their son to answering his own question.

Communication Breakdown 2: Family B

The next conversational interaction is with Family B, which begins in a similar way to Communication Breakdown 1 in Family E. One family member first consults other family members for information, but when other family members are not able to provide the information, the family turns to technology for the answer.

(Mother): Hey [daughter's name]! Hey [father's name]!
 (pause) South Korea (pause) what part of Asia? Southeast or East Asia?
 (indistinct family member response)
 (Mother): From America. What you . . .
 (indistinct family member response)
 (Mother): Yeah.
 (Father): Google is South Korea East Asia or South?
 (Mother): Alexa! South Korea is East Asian or Southeast Asian?
 (Alexa): Sorry, I'm not sure.
 (Daughter): Alexa. Are Koreans Southeast Asian?
 (Alexa beeps)
 (pause in family conversation)
 (Mother): Try again [daughter's name].
 (Daughter, with increased volume): Alexa! Are Koreans Southeast Asian?
 (Alexa): I'm not quite sure how to help you with that.
 (Daughter): Hmmm...
 (Mother): Alexa. Who is Eastern East Asia? What country is East Asia?
 (Alexa): There are no UN recognized countries in Eastern Europe.
 (Daughter): Augh..
 (Father, with raised voice): Alexa. Please define East Asian countries.

(Alexa): This might answer your question. There are five UN recognized countries in Eastern Asia: China, Japan, South Korea, North Korea, and Mongolia. Did that answer your question?

(Mother): Yes.

(Alexa): Thanks for your feedback.

This interaction illustrates multiple types of communication repairs that families in our study used: syntax adjustments, semantics adjustments, changes in volume, changes in prosody, and collaboration with family members. To begin, the mother consults other family members for information, and when they are not able to provide that information, the father suggests the family's go-to information resource: Google. By doing this, the father immediately provides discourse scaffolding and shifts the family's focus to include media. The mother decides to try Alexa first, essentially shifting the family's media attention from Google as a resource to Alexa. The mother's question to Alexa closely follows the syntactical model suggested by the father for the Google search, with an expansion to clarify the last key term "*Southeast Asian*" vs "*South*." This form of question is essentially an either/or type of question, in which the requester provides two options and is looking for a response that fits one of those options. Alexa's response, instead, is a neutral response, "*Sorry, I'm not sure*," which does not provide her communication partner with any specific information on the cause of the communication breakdown.

The daughter, originally asked as a primary information resource, actively takes part in the conversation with Alexa and takes a turn at repairing the communication breakdown. She uses her mother's query as a foundation for her interaction. The daughter employs contraction by making a syntactical change: deleting the words in the query and thus changing the nature of the question to a yes/no question ("*Are Koreans Southeast Asian?*"). Alexa's response to this question is an extreme version of a neutral response, that of a beep and silence. The family conversation briefly comes to a halt with this response from Alexa, causing a second, and perhaps more dramatic communication breakdown.

After the brief silence, the mother encourages her daughter to try again through direct instruction. At this point, both mother and daughter are actively engaged in jointly repairing the communication breakdown. The daughter employs repetition with increased volume in this conversational repair attempt (the third total conversational repair attempt for the family as a unit). Alexa responds with another neutral response, "*I'm not quite sure how to help you with that*."

At this point, the mother resumes her original role in communicating with Alexa and attempts the fourth repair. She modifies her daughter's interaction through a combination of expansion and contraction. Utilizing a dramatic syntactic change, she breaks up her original either/or question into two closed ended questions, asking both "who" and "what" questions.

By now, we have seen the original question modified five times by three family members. Each time the question is modified in relation to the prior family member's version of the question. Up to this point, Alexa has provided neutral responses, which resulted in the family collaborating to make a wide variety of repair types. Without any specific responses or actions from Alexa, the family did not have any signposts to guide their repair strategies.

In response to the mother's two closed ended questions, Alexa acts on a misunderstanding. After the fourth repair attempt is unsuccessful, the daughter makes a noise of frustration and at this point the father actively joins the interaction with Alexa for the first time. His interaction indicates that he has been listening to the prior communication attempts between his family and Alexa but has waited to take his turn at communication repair. The father's communication repair employs another syntactic change, as well as increasing his volume. He shifts the question to a specific, closed-ended request: "*Please define East Asian countries*." He also redirects Alexa from an off-topic response (countries in Eastern Europe) back to the specific topic the family is interested in.

All three family members collaborated to achieve a successful communication interaction with Alexa. Each family member listened to the interaction between Alexa and other family members and adjusted their communication interaction based on the prior interactions. At the end, after providing the desired information, Alexa instigates turn-taking with the family by asking them a question "*did that answer your question?*" The mother, the initiator of the communication interaction, responds to Alexa's question. Alexa responds back to the mother with a statement that concludes the communication interaction on this topic.

Communication Breakdown 3: Family D

Our final collaborative communication repair example is from one of our two bilingual families. We note that this final communication breakdown example is representative of two key themes found across our participant families: children's speech and language development impacting communication with Alexa (captured in multiple

instances, but primarily with monolingual children) and Alexa struggling to understand bilingual individuals (also captured in multiple instances, but primarily with adults). Both of these themes were reinforced by family comments during our final interviews. The communication breakdown described below is the only communication breakdown that combined these two themes in one example.

This conversational interaction is from Family D, a family that speaks both English and Spanish in their home. In this example, the family has set up their Echo Dot for the first time and have just completed the guided setup through the Alexa app with the researchers in the room. The mother encourages their 5-year-old to ask a question, but the 5-year-old has trouble with a consistent production of the word “Alexa.”

(Mother): It’s A (pause) lexa (emphasis on the last two syllables)
 (Child): uh (slight pause) leh (pause) ska. Is it going to rain for a little bit or is it going to be sunny for a little bit. (said quickly and quietly), (pause) or both. (no rising intonation at end to indicate a question)
 (pause and no response from Alexa)
 (Researcher): That was a good question but it might have been a little too long, too many questions I heard.
 (Mother): Let’s see. (pause) Alexa, is it going to be rainy all day or sunny all day or . . . (recording cut off).
 (recording resumes with child giggling)
 (Mother): Alexa, what did I ask you?
 (Alexa): Sorry, I’m not sure
 (everyone laughs)
 (Mother): Alexa, I asked if it was going to rain all day.
 (Alexa): Probably not. Each day of the next 7 days . . . has at most a 30% chance of rain.
 (child laughs)

Joint engagement with Alexa is clear for the entirety of the communication interaction, evidenced by the communication partners’ reactions to all interactions with Alexa. Initially, the child’s mother provides direct instruction to her child on improving the production of the word “Alexa.” We found that many of the young children in our study struggled with the speech production necessary to make themselves consistently understood by Alexa. Developmentally, children at age 5 can be expected to have difficulty with the consistent production of some sounds, including the /l/ sound and /s/ sound, which are both part of the word “Alexa” [44]. Typically developing young bilingual children can be expected to make sound substitutions from one language to another [15]. We see evidence of some of these

developmental speech patterns with the child in Family D in our conversational sample described earlier.

Another theme we find in this example is the linguistic construction of the question the child asks Alexa: “*Is it going to rain for a little bit or is it going to be sunny for a little bit, or both?*” The syntax of the question is complex, asking Alexa three different questions about the weather grouped into one. We also see that the child’s use of language does not match Alexa’s abilities. This is the first time this child has used a digital home assistant, and the child is most likely unaware of the need to code-switch. Therefore, the researcher provides direct instruction and suggests the child might have asked too many questions at once.

The mother then modifies her child’s question, by adjusting the syntax and semantics, to attempt her communication repair. However, Alexa does not respond. The mother then makes another attempt at communication repair by asking Alexa, “*what did I ask you?*” Alexa provides a neutral clarification response. Joint engagement in the interaction is demonstrated by everyone laughing at Alexa’s response.

In the final communication repair, the mother significantly modifies both the syntax and semantics of the question to a contracted yes/no question. Alexa then responds appropriately to the question, which ends that specific communication interaction.

Pragmatics: Code Switching with Alexa

Although we find that most family members in our study are able to adapt communication interactions with Alexa, as evidenced through the multiple types of repair strategies used, not everyone was successful in this process. In particular, children under the age of 5 tended to struggle more than older children and adults. For example, the parents of Family I report that one of their younger children (age 4) would walk up to the Echo Dot and begin talking to Alexa without first addressing her by name. The child would use long sentences and often change topic before Alexa responded. His parents felt that his communication with Alexa was similar to his communication with other people. The parents believed that their son expected Alexa to respond to him like other people do, and they reported that he became frustrated when she did not. In contrast, their older child (age 6) would say Alexa’s name and then pause before continuing his communication interaction with her.

It was not only young children that had a difficult time code switching. We find that Alexa did not code switch with family members to craft her communication

interactions to the needs and abilities of the person she was addressing, as evidenced most clearly through the jokes she told. For example, the child in Family J asked Alexa to tell jokes throughout the four-week deployment. She responded with a variety of jokes, such as: “*A man walks into a bar. Crank. It was a heavy metal bar.*” During our final interview with Family J, we asked the mother and child if the jokes Alexa told were appropriate for the child (age 4). The mother replied “*Yeah, there were a couple that [child’s name] chuckled at but I don’t think she (pause and mother directs question to child) ‘do you know why the jokes were funny? Did they make sense to you?’*” The child’s response was “*No, I don’t know why.*”

Alexa’s lack of code switching is not only evident with children. We find that Alexa is unable to code switch with multilingual users. For example, when the mother in Family D says, “*Alexa, play song by Marco Antonio Solis,*” producing the singer’s name using Spanish phonological processes instead of an English phonetic version of the singer’s name. Alexa’s response is “*‘Unbelieve,’ by Andy Groun.*” In this case, Alexa misheard or misunderstood the request because of her inability to code switch to recognize that the requester was using the correct pronunciation for words in another language.

We can think of Alexa as being very literal with communication interactions, which implies a lack of understanding of the social nuances and uses of language, essentially, a deficit in pragmatic skills. An excellent example of Alexa’s poor pragmatic skills is with Family A:

(Mother): Alexa, add 157 (*slight pause*)
 (Alexa): I added 157 to your shopping list.
 (Mother): Alexa, calculate 157 plus 50 (*pause*)
 (Alexa): Sorry, I didn’t catch that.

Alexa did not understand that the pause in this case was in anticipation of continuing the request. In addition, Alexa did not understand that the semantic modification, from “*add*” to “*calculate*” was intended to provide a definition of the earlier request of the word “*add*.” Multiple meanings of words are a challenge for Alexa, and for any communication partner that does not understand the context of the communication interaction. Finally, Alexa failed to recognize the contextual cue that “*157*” was a number and therefore the correct interpretation of “*add*” is to calculate. A final example of Alexa’s inability to code switch is when the child in Family J asks Alexa to count to 10 in Spanish. Alexa responds by translating the word “*count*” from English to Spanish.

5 DISCUSSION

Currently, the burden of repairing communication breakdowns with Alexa lies with family members. Repairing communication breakdowns with Alexa can take a great deal of effort on the part of multiple family members, and in some cases, might lead to abandonment. Our findings reinforce Suchman’s suggestion from 1987: designers of machines should focus on designing for collaboration between the user and the machine to achieve mutual understanding [39]. To alleviate the burden on family members, and to increase chances of successful conversational repairs, we recommend that digital home assistants are designed with a focus on improving communication repair strategies.

Digital Code Switching for Human Interaction

Understanding the communicative context and the communication partners involved in a conversation incorporates the field of pragmatics [35,36]. To improve communication repair strategies, knowledge of the context and the communication partner is extremely helpful, allowing digital home assistants to artificially code switch as needed. We acknowledge that Alexa (the Echo Dot) is a machine, and does not have pragmatic skills; however, it may be possible to build on the concept of code switching to design more appropriate interactions between digital home assistants and humans. Designers who are mindful of the varied skill levels and communication styles of family members can improve digital assistants’ abilities to repair communication breakdowns. Younger children have difficulty with certain sounds and sound combinations [44], and they tend to use repetition as an initial communication repair strategy [25]. If designers imbued voice assistants with the skills to identify that a communication partner is a young child, and therefore is likely to use repetition as a repair strategy and likely to have imprecise speech, a digital home assistant can provide more specific communication feedback to collaborate on the communication repair. One way in which the specific feedback can be helpful is through the form of discourse scaffolding.

Discourse Scaffolding and Specific Clarification Responses

Our findings demonstrate how parents use discourse scaffolding to direct their child’s communication interactions. With a designed ability to code-switch, Alexa can do the same, based on recognizing cues from the human communication partner. Developing discourse scaffolds is an expansion of some basic functions that

Alexa already has, such as performing specific clarification responses like, “*Do you mean 10:45 in the morning or the evening?*” By developing the ability to code switch, Alexa can apply directive discourse scaffolding to a greater variety of communication interactions, based on her understanding of the context and her communication partner. For example: Alexa could identify repetition and imprecise speech and respond with, “*I’m not quite sure what you’re asking. Can you tell me just the key word of what you want?*”

Other work in the area of conversational interfaces has made similar suggestions for conversational agents to provide cues and prompts to collaborate with their human conversational partners [9,12,19,34,43]. Here, we further suggest that the constructs of discourse scaffolding and code switching could be productively applied to the design of digital assistants. Discourse scaffolding can be applied to any conversational partner (adult or child), and through code switching, customized to provide the level of support that is appropriate for the specific conversational partner. Our close analysis of conversational breakdowns and repairs with family members provides an initial guide for developers to design conversational strategies which incorporate code switching and discourse scaffolding.

We realize that the application of code switching may inadvertently create new miscommunications between Alexa and her communication partners. Alexa might misidentify an adult using repetition and talking while chewing gum as following the same patterns as a young child. However, the use of specific clarification responses provides more signposts for communication repairs than neutral clarification responses. If Alexa makes a mistake, it is easier for her human communication partner to identify the cause and make the repair with specific responses. Ideally, the responses from her human communication partner allow Alexa to learn her communication partner’s speech patterns over time, making code switching easier, and increasing the accuracy of communication responses.

Promoting Family Engagement and Communication

Based on our findings, we see how the concept of joint media engagement [40] is useful for designers to consider when creating applications and skills for digital home assistants. Due to the conversational nature of the interaction, other people in the immediate environment listen to and engage in communication with digital home assistants when there is a communication breakdown, both at home and in other social settings [33,34]. Designers of digital home assistants can capitalize on this

knowledge and engage families in positive, joint interactive experiences, even when only one family member initially engages with the digital home assistant. We saw from our participant families that some family members did not engage immediately with Alexa until other family members had engaged in multiple interaction attempts. For example, when designing an Alexa skill for multiple players, such as a game or competition, family members might be encouraged to join in midway through: “*Is there anyone else who can jump in and help?*”

Utilizing the concept of joint media engagement, digital home assistants can simultaneously build their own pragmatic skills while assisting family members with their own conversational skills. If the digital home assistant is transparent about the nature of the communication breakdown, it can increase family members’ own awareness of their communication skills. In this way, the digital home assistant can model meta-pragmatic skills, in which people talk about their own social use of language and their understanding of the social use of language [41]. Digital home assistants have the potential to become communication facilitation tools for communication both between family members and between family members and the digital home assistant.

Based on our analysis of diverse families’ communication breakdowns and repairs with Alexa, we find many opportunities to support the growth of Alexa’s communication skills. We particularly want to emphasize the potential for digital home assistants to assist their human communication partners in repairing communication breakdowns. Our interactions with technology have radically changed since Suchman’s work in 1987, however, the hurdle of overcoming communication breakdowns remains the same [39]. Voice interfaces provide a unique opportunity for technology to take on the burden of developing a shared understanding with their human communication partners.

Limitations

The corpus of communication interactions analyzed for this study is a subset of the total communication interactions captured and likely excludes additional examples of communication breakdowns and repairs. Our findings are also limited in that some families expressed an awareness of needing to “use” the Echo Dot during the study, and this may have perpetuated their attempts to repair communication breakdowns with Alexa. To mitigate that effect, our corpus includes communication breakdown segments that occurred throughout the four-week study period. Our sample is also limited to 10

families in one area of the United States. Future work remains to conduct a larger scale analysis of families in different regions with a broader range of language and dialectical differences.

6 CONCLUSION

The burden of repairing communication breakdowns between humans and technology continues to rest with humans. Our findings highlight how families collaborate in a variety of ways to repair communication breakdowns with digital home assistants, often requiring multiple repair attempts with more than one family member trying to successfully communicate with Alexa (the conversational name used with the Echo Dot). While Alexa makes some attempts to repair communication breakdowns through specific clarification responses, neutral clarification responses are far more common, which do not aid in helping the human communication partner repair the breakdown.

We turn back to 1987 and Suchman’s guidance that designers focus on improving how technology can promote shared understanding with their human communication partners. Rather than trying to anticipate a user’s needs, the digital home assistant can promote conversational collaboration through improvements in repairing communication breakdowns. To this end, we suggest that designers incorporate specific processes which promote the artificial pragmatic skills of digital home assistants. Improving technology’s ability to identify their human communication partners and to provide specific clarification responses will ultimately improve the ability for humans and machines to collaborate towards shared understanding in their conversational interactions.

ACKNOWLEDGMENTS

This research was funded, in part, by a grant from Mozilla. We also want to acknowledge our appreciation of the speech-language pathologists who we consulted for validity checks of our coding for communication repair types: Teresa Fleck, a practicing SLP of 20 years, and Victoria Lai, an SLP and researcher. We also appreciate the transcription and administration support of Bella Chiu, Yini Guan, Leanne Liu, Sharon Heung, Ashley Boone and Kate Yen. Finally, we sincerely thank the participant families, without whom this research would not be possible.

REFERENCES

- [1] Linda Bell and Joakim Gustafson. Interaction with an Animated Agent in a Spoken Dialogue System. Retrieved September 13, 2018 from https://www.speech.kth.se/august/eur99_augdata.html
- [2] D. Bohus and A.I. Rudnicky. 2001. Modeling the cost of misunderstanding errors in the CMU Communicator dialog system. In *IEEE Workshop on Automatic Speech Recognition and Understanding*, 2001. ASRU '01., 252–255. DOI:<https://doi.org/10.1109/ASRU.2001.1034635>
- [3] Dan Bohus and Eric Horvitz. 2009. Open-world dialog: Challenges, directions, and prototype. In *Proceedings of IJCAI'2009 Workshop on Knowledge and Reasoning in Practical Dialogue Systems*.
- [4] Dan Bohus and Eric Horvitz. 2011. Multiparty Turn Taking in Situated Dialog: Study, Lessons, and Directions. In *Proceedings of the SIGDIAL 2011 Conference (SIGDIAL '11)*, 98–109. Retrieved from <http://dl.acm.org/citation.cfm?id=2132890.2132903>
- [5] Holly P. Branigan, Martin J. Pickering, Jamie Pearson, and Janet F. McLean. 2010. Linguistic alignment between people and computers. *J. Pragmat.* 42, 9 (September 2010), 2355–2368. DOI:<https://doi.org/10.1016/j.pragma.2009.12.012>
- [6] Bonnie Brinton, Martin Fujiki, Diane Frome Loeb, and Erika Winkler. 1986. Development of conversational repair strategies in response to requests for clarification. *J. Speech Hear. Res.* 29, 1 (March 1986), 75–81. DOI:<https://doi.org/10.1044/jshr.2901.75>
- [7] Antony Bryant and Kathy Charmaz. 2007. *The SAGE handbook of grounded theory*. London: SAGE Publications, London.
- [8] Li-Rong Lilly Cheng. 2007. Codes and Contexts: Exploring Linguistic, Cultural, and Social Intelligence. *ASHA Lead.* 12, 7 (May 2007), 8–33. DOI:<https://doi.org/10.1044/leader.FTR2.12072007.8>
- [9] 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 (IDC '18)*, 337–348. DOI:<https://doi.org/10.1145/3202185.3202749>
- [10] Benjamin R. Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. 2017. “What Can I Help You with?": Infrequent Users’ Experiences of Intelligent Personal Assistants. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*, 43:1–43:12. DOI:<https://doi.org/10.1145/3098279.3098539>
- [11] John W. Creswell. 2018. *Qualitative inquiry & research design: choosing among five approaches* (Fourth ed. ed.). SAGE, Thousand Oaks, CA.
- [12] Stefania Druga, Randi Williams, Cynthia Breazeal, and Mitchel Resnick. 2017. “Hey Google is It OK if I Eat You?": Initial Explorations in Child-Agent Interaction. In *Proceedings of the 2017 Conference on Interaction Design and Children (IDC '17)*, 595–600. DOI:<https://doi.org/10.1145/3078072.3084330>
- [13] Paula B. García, Lori Leibold, Emily Buss, Lauren Calandrucio, and Barbara Rodriguez. 2018. Code-Switching in Highly Proficient Spanish/English Bilingual Adults: Impact on Masked Word Recognition. *J. Speech Lang. Hear. Res.* (August 2018), 1–11. DOI:https://doi.org/10.1044/2018_JSLHR-H-17-0399
- [14] Catherine Garvey. 1975. Requests and responses in children’s speech”. *J. Child Lang.* 2, 1 (April 1975), 41–63. DOI:<https://doi.org/10.1017/S03050009000088X>
- [15] Christina E. Gildersleeve-Neumann, Ellen S. Kester, Barbara L. Davis, and Elizabeth D. Peña. 2008. English Speech Sound Development in Preschool-Aged Children From Bilingual English–Spanish Environments. *Lang. Speech Hear. Serv. Sch.* 39, 3 (July 2008), 314–328. DOI:[https://doi.org/10.1044/0161-1461\(2008/030\)](https://doi.org/10.1044/0161-1461(2008/030))
- [16] Alexis Hiniker, Bongshin Lee, Kiley Sobel, and Eun Kyoung Choe. 2017. Plan & Play: Supporting Intentional Media Use in Early Childhood. In *Proceedings of the 2017 Conference on Interaction Design and Children (IDC '17)*, 85–95. DOI:<https://doi.org/10.1145/3078072.3079752>
- [17] Yvette D. Hyter. 2007. Pragmatic Language Assessment: A Pragmatics-As-Social Practice Model. *Top. Lang. Disord.* 27, 2 (April 2007), 128–145. DOI:<https://doi.org/10.1097/01.TLD.0000269929.41751.6b>
- [18] Sharon L. James. 1990. *Normal Language Acquisition*. Allyn & Bacon, Incorporated.
- [19] Jiepu Jiang, Wei Jeng, and Daqing He. 2013. How Do Users Respond to Voice Input Errors?: Lexical and Phonetic Query Reformulation in Voice Search. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '13)*, 143–152. DOI:<https://doi.org/10.1145/2484028.2484092>
- [20] Alan Kennedy, Alan Wilkes, Leona Elder, and Wayne S. Murray. 1988. Dialogue with machines. *Cognition* 30, 1 (October 1988), 37–72. DOI:[https://doi.org/10.1016/0010-0277\(88\)90003-0](https://doi.org/10.1016/0010-0277(88)90003-0)
- [21] Margaret Lahey. 1988. *Language Disorders and Language Development* (1 edition ed.). Pearson, New York: London.
- [22] M. Lohse, K. J. Rohlfing, B. Wrede, and G. Sagerer. 2008. “Try something else!” – When users change their discursive behavior in human-robot interaction. In *2008 IEEE International Conference on Robotics and Automation*, 3481–3486. DOI:<https://doi.org/10.1109/ROBOT.2008.4543743>
- [23] Ewa Luger and Abigail Sellen. 2016. “Like Having a Really Bad PA”: The Gulf Between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*, 5286–5297. DOI:<https://doi.org/10.1145/2858036.2858288>
- [24] Barbara G. MacLachlan and Robin S. Chapman. 1988. Communication Breakdowns in Normal and Language Learning-Disabled Children’s

- Conversation and Narration. *J. Speech Hear. Disord.* 53, (1988). DOI:<https://doi.org/10.1044/jshd.5301.02>
- [25] Tova Most. 2002. The Use of Repair Strategies by Children With and Without Hearing Impairment. *Lang. Speech Hear. Serv. Sch.* 33, 2 (April 2002), 112–123. DOI:[https://doi.org/10.1044/0161-1461\(2002/009\)](https://doi.org/10.1044/0161-1461(2002/009))
- [26] Clifford I. Nass, Youngme Moon, John Morkes, Eun-Young Kim, and B. J. Fogg. 1997. Human Values and the Design of Computer Technology. In Batya Friedman (ed.), *Center for the Study of Language and Information*, Stanford, CA, USA, 137–162. Retrieved August 31, 2018 from <http://dl.acm.org/citation.cfm?id=278213.278222>
- [27] M. A. Norman and P. J. Thomas. 1991. Informing HCI design through conversation analysis. *Int. J. Man-Mach. Stud.* 35, 2 (August 1991), 235–250. DOI:[https://doi.org/10.1016/S0020-7373\(05\)80150-6](https://doi.org/10.1016/S0020-7373(05)80150-6)
- [28] Sharon Oviatt, Jon Bernard, and Gina-Anne Levow. 1998. Linguistic adaptations during spoken and multimodal error resolution. *Lang. Speech* 41, 3/4 (July 1998), 419–442.
- [29] Michael Quinn Patton. 2015. *Qualitative research & evaluation methods: integrating theory and practice* (Fourth ed. ed.). SAGE Publications, Inc, Thousand Oaks, California.
- [30] Jamie Pearson, Jiang Hu, Holly P. Branigan, Martin J. Pickering, and Clifford I. Nass. 2006. Adaptive Language Behavior in HCI: How Expectations and Beliefs About a System Affect Users' Word Choice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*, 1177–1180. DOI:<https://doi.org/10.1145/1124772.1124948>
- [31] Hannah R.M. Pelikan and Mathias Broth. 2016. Why That Nao?: How Humans Adapt to a Conventional Humanoid Robot in Taking Turns-at-Talk. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, 4921–4932. DOI:<https://doi.org/10.1145/2858036.2858478>
- [32] Sarah Perez. Voice-enabled smart speakers to reach 55% of U.S. households by 2022, says report. *TechCrunch*. Retrieved September 5, 2018 from <http://social.techcrunch.com/2017/11/08/voice-enabled-smart-speakers-to-reach-55-of-u-s-households-by-2022-says-report/>
- [33] 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 (CHI '18)*, 640:1–640:12. DOI:<https://doi.org/10.1145/3173574.3174214>
- [34] Martin Porcheron, Joel E. Fischer, and Sarah Sharples. 2017. “Do Animals Have Accents?”: Talking with Agents in Multi-Party Conversation. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*, 207–219. DOI:<https://doi.org/10.1145/2998181.2998298>
- [35] Carol A. Prutting. 1982. Pragmatics as Social Competence. *J. Speech Hear. Disord.* 47, 2 (May 1982), 123–134. DOI:<https://doi.org/10.1044/jshd.4702.123>
- [36] Kenyatta O. Rivers, Yvette D. Hyter, and Glenda DeJarnette. 2012. Parsing Pragmatics. *ASHA Lead.* 17, 13 (October 2012), 14–17. DOI:<https://doi.org/10.1044/leader.FTR1.17132012.14>
- [37] Alex Sciuto, Armita Saini, Jodi Forlizzi, and Jason I. Hong. 2018. “Hey Alexa, What’s Up?”: A Mixed-Methods Studies of In-Home Conversational Agent Usage. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*, 857–868. DOI:<https://doi.org/10.1145/3196709.3196772>
- [38] Tina Sharpe. 2006. “Unpacking” Scaffolding: Identifying Discourse and Multimodal Strategies that Support Learning. *Lang. Educ. Int. J.* 20, 3 (May 2006), 211–231.
- [39] Lucy A Suchman. 1987. *Plans and situated actions: The problem of human-machine communication*. Cambridge university press.
- [40] Lori Takeuchi and R Stevens. 2011. The new coviewing: Designing for learning through joint media engagement.
- [41] Geraldine P. Wallach and Katharine G. Butler. 1994. *Language Learning Disabilities in School-Age Children and Adolescents: Some Principles and Applications* (1 edition ed.). Pearson, New York: Toronto: New York.
- [42] Akiko Yamazaki, Keiichi Yamazaki, Matthew Burdelski, Yoshinori Kuno, and Mihoko Fukushima. 2010. Coordination of verbal and non-verbal actions in human-robot interaction at museums and exhibitions. *J. Pragmat.* 42, 9 (September 2010), 2398–2414. DOI:<https://doi.org/10.1016/j.pragma.2009.12.023>
- [43] Svetlana Yarosh, Stryker Thompson, Kathleen Watson, Alice Chase, Ashwin Senthilkumar, Ye Yuan, and A. J. Bernheim Brush. 2018. Children Asking Questions: Speech Interface Reformulations and Personification Preferences. In *Proceedings of the 17th ACM Conference on Interaction Design and Children (IDC '18)*, 300–312. DOI:<https://doi.org/10.1145/3202185.3202207>
- [44] Four to Five Years Old. *American Speech-Language-Hearing Association*. Retrieved August 20, 2018 from <https://www.asha.org/public/speech/development/45/>