

BIELEFELD UNIVERSITY

SEMINAR TITLE: MISCOMMUNICATION AND REPAIR IN VOICE ASSISTANTS

Communication Breakdowns Between Families and Alexa

Author:

Syed Zain Ali

Data Science's Student

4373556

Date: October 7, 2024

Contents

Affidavit

Abstract

1 Introduction	1
2 Background/Related Work	2
3 Analyzing Communication Breakdowns Between Families and Alexa	3
3.1 Methodology of Compared Studies	3
3.2 Types of Communication Breakdowns	4
3.2.1 Communication Breakdown between Families & Alexa	4
3.2.2 Alexa, You're Really Stupid	7
3.3 Comparison of Findings	10
3.4 Conclusion	10
4 Repair Strategies of Families	11
4.1 Communication Breakdown between Families and Alexa	11
4.2 Alexa, you're really stupid	13
4.3 The 4A's	14
4.4 Miscommunication Detection and Recovery in Situated Human-Robot Dialogue	15
4.5 Exploring humor as a repair strategy during communication breakdowns with	
voice assistants	15
4.6 Family Negotiation in Joint Media Engagement with Creative Computing . .	16
4.7 Comparison of findings	17
5 Conclusion	18
References	20
Abbreviations	21
List of Tables	22
List Of Figures	22

Declaration of Consent (Affidavit)

“Hereby I declare that I wrote this essay independently and that I prepared all graphical representations and tables included independently. I did not use any sources other than those indicated and always marked those parts of the essay, which are taken from other works, as borrowed from a specific clearly-stated source, including tables and illustrations. This essay was personally written by me and not generated by any largescale language model, such as Chat-GPT or similar.”

Syed Zain Ali

Bielefeld, October 7, 2024

A handwritten signature in black ink, appearing to read 'Zain', with a horizontal line underneath it.

Syed Zain Ali

Abstract

In this essay, I have described different research papers and analysed family's communication with digital voice assistants like Alexa. The purpose of this analysis is to understand communication breakdowns, their reasons, and how these breakdowns are handled by family members. I studied mainly 2 papers 'Communication Breakdown between families & Alexa' [1] and 'Alexa you're really stupid' [2] to understand reasons for breakdowns and mistakes in conversation and their recovery. Furthermore, I studied 5 different papers to describe breakdown repairs along with the first two papers and analysed common methodologies by families and which strategies were useful and which were not. One important aspect that is described along with other problems is user expectation and satisfaction and its relationship with device dropout. Designers can adapt these methodologies for better efficiency on the basis of these analyses, which will improve error handling and increase user productivity, satisfaction level, and efficiency of digital voice assistants.

1 Introduction

Human communication began around 100,000 BCE, and the invention of the first writing system occurred approximately at the start of the Bronze Age [3]. Later on, after many evolutions' communication shaped to symbols and then languages. However, communication doesn't start earlier, directly from words. Communication is an important aspect of our daily lives. It is considered to be one of the key components of any individual or organisation's success. Similarly, miscommunication can lead to negative or sometimes severe consequences. Therefore, it's important to deliver effective words and make sure others understand you as well. Many scholars have tried to give a definitive meaning to communication. In the view of the Association of Perioperative Practice, communication originated from the Latin word *communis*. Lunenberg (2010) agrees with them that communication has its root in Latin from the word *communis*. It is said that communication is the art and technique of using words effectively to impart information or ideas [4]. Now in the modern era, communication is not only limited to humans but extended to machines as well, gaining more and more popularity, especially due to the invention of digital home assistants like Alexa, smartphones, chatbots, and information systems. For humans, it's easy to identify miscommunication and solve errors, but in machines, this is quite complicated, and machines interpret what they have learned so far. If machines misinterpret something, either they'll be unable to give the right answer or will be unable to understand something completely. Miscommunication is defined by Wadensjö (1998, p. 198) as 'lack of fit between the sense aimed at by one interlocutor and what is displayed by another as the sense made of the current message'. [5] In this essay, I have explained communication breakdowns between families and voice assistants. I focused on the explanation and reasoning of communication breakdowns and their repair in families. My research questions are:

Research Questions

- What types of communication breakdowns occur?
- How do families repair communication breakdowns with Alexa?

Furthermore, I've explored what kinds of strategies people use when breakdowns occur and

how they repair breakdowns. To elaborate my problem statement, I explain research papers specifically Communication Breakdowns Between Families and Alexa, Exploring Humor as a Repair Strategy During Communication Breakdowns with Voice Assistants, and The 4A's. From the first paper, I have explained how communication breakdowns occur and of which type. How do families handle these breakdowns? Either their problem is solved or not? Moreover, from the second and third papers, I have explained strategies for how families can avoid communication errors or breakdowns or how they can handle if a breakdown occurs. However, before diving deeply, I give technical details thoroughly and most important strategies and methodologies. It includes the involvement of speech-language pathologists to analyze breakdowns, joint media engagement between families to handle errors, discourse scaffolding, and briefly explaining human-computer interaction. I explained family negotiation in joint media engagement with creative computing to tackle communication errors. At the end, I give an analysis of different papers that explains the reasons for breakdowns and strategies adapted to overcome the breakdown. Then I give a conclusion involving my own expertise and opinion about reasons and strategies that can be adapted to explain how we can avoid and handle communication breakdowns in digital voice assistants.

2 Background/Related Work

The field of human communication and communication breakdowns includes pragmatic topics related to communication like turn-taking, topic maintenance, and code switching, which take many forms. We'll discuss further how these methods are used to handle breakdowns and whether they were successful or not. In the field of human-computer interaction, a lot of work has been done. In 1987, Suchman framed acts of human-machine interaction as a dialogue between communication partners [6]. Furthermore, HCI researchers explain the role of parents and joint media in broad terms of negotiation instead of the parent's perspective. A broad analysis of advocacy for control to identify negotiations in families is explained in the paper Family Negotiation in Joint Media Engagement with Creative Computing. [7] Matthew Marge and Alexander I. Rudnicky explain repair strategies in their research paper Miscommunication Detection and Recovery in Situated Human-Robot Dialogue, published on 17-Feb-2019. [8] In this paper, they have explained the nearest neighbor algorithm to solve a

problem. A robot interacts with a previous problem's history to determine how it has resolved a particular situation. I found in the paper 'Communication Breakdowns between Families & Alexa' in conclusion section that they suggested designers should incorporate artificial pragmatic skills in digital devices, which will enhance mutual understanding between humans and machines. Common issues such as grammatical errors, semantic misunderstandings, and the misinterpretation of commands by home assistants can lead to user dissatisfaction, which may ultimately result in users abandoning these devices. As mentioned in the discussion section of 'Alexa, You're Really Stupid,' this often happens when the devices fail to meet user expectations. Most of the voice assistants (VA) users want them to be more than just a conversational partner and help in chaos, hence making their life easy.

3 Analyzing Communication Breakdowns Between Families and Alexa

3.1 Methodology of Compared Studies

I will explain the above-mentioned research papers, and for communication breakdown analysis, I have explained two research papers, and after finding the breakdowns and types of breakdowns in each paper, I have explained both in detail with reasons and similarity of breakdowns that are common in both papers. The studies analysed in this essay used various methods to gather data on communication breakdowns between families and digital home assistants. Research paper "Communication breakdowns between families & Alexa" recorded interactions of ten families, each with at least one child, to explore the types of breakdowns and the strategies used to repair them. The second research paper, "Alexa, You're Stupid," explains the study conducted a five-week longitudinal analysis, measuring user satisfaction, repair strategies, and the emotional responses of children and adults. These methodologies provide a comprehensive understanding of the common issues and potential solutions in human-Alexa interactions. Both research papers have some limitations like small sample size. CMB between families & Alexa describes 10 families over a short four-week period, which may not capture long-term behaviors, and they focused exclusively on Amazon Alexa, limiting the applicability to other devices. Additionally, technical issues in record-

ing interactions led to incomplete data collection. Similarly, second research paper [2] ‘Alexa You’re really Stupid’, faced limitations with its small sample of nine families and the fact that the original recordings, conducted in German, had to be translated into English, potentially introducing translation errors. Moreover, this research relied on log files and self-reported data rather than real-time observations, which could affect the depth and accuracy of the data collected. Both studies also focused solely on Amazon Alexa, restricting their findings to a single device.

3.2 Types of Communication Breakdowns

In this section, I am going to explain key findings of both research papers one by one to explore communication errors, common mistakes, and reasons for these errors.

3.2.1 Communication Breakdown between Families & Alexa

First of all, I have explained types of communication breakdowns and analysed their reasons. At first, I found some common breakdowns and causes in the research paper Communication Breakdown between Families & Alexa. In this paper, researchers have recorded data of 10 families, and each family has a child, and two of the families were bilingual. From this paper, I will explain three communication examples from this paper to explain breakdown. It’s not convenient for humans to talk to conversational agents, as they talk to humans in spite of having a conversational interface. Speech language pathologists analyze the recordings, use speech-language constructs, and validate types of communication repairs. It is said almost all families have faced the problem. What I found from this paper is that most of the problems were because of a lack of communication skills in families or interference from something like noise. The breakdowns happen because of poor pragmatic skills semantically incorrect sentence or word, synthetically incorrect commands, inability of user to construct words which are linguistically correct, or sometimes the inability of device to understand semantics modification. Most of the communication breakdowns occur either because of one of the problems described above or the combination of these problems.

Response type	Definition	Example
Acting on Misunderstanding (AoM)	Performing an action or providing a response based on misheard or misunderstood input	<i>"Alexa, what should we do this night?"</i> Alexa responds with a definition of "this night."
Neutral Clarification Response (NR)	Providing an indication that the communication partner's interaction was unclear	<i>"Sorry, I don't know that."</i> <i>"Sorry, I'm not sure."</i> <i>"I'm not quite sure how to help you with that."</i>
Specific Clarification Response (SR)	Providing specific information to request clarification	<i>"Did you mean [person's name]?"</i> <i>"Is that 10:45 in the morning or evening?"</i>

Table 1: Signals from Alexa that a communication breakdown has occurred (Communication Breakdowns between Families & Alexa, page no.6, Table 4)

Furthermore, communication breakdowns are not only from humans but can happen also from machines. For example, in this paper I got 3 types of responses from Alexa that indicate a breakdown has occurred that are described in table1. which indicates that designers's algorithms are at least good enough that they can capture the problem and are not completely dull. I'll illustrate one example from this paper where a 5-year-old child try to interact with Alexa, and this family is bilingual. Below is the complete dialogue taken from research paper "Communication breakdowns between Families & Alexa" page 9, Family D.

(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)

Figure 1: Image taken from research paper [1] page 9, Family D conversation

After reviewing and analyzing the conversation, I found that because of the child's bilingualism and premature age, the child is inconsistent to pronounce the word 'Alexa', and his mother teaches her to pronounce it correctly. As we can notice, Child asks a question to Alexa in an irregular manner and infers, "is it going to rain for a little bit or is it going to be sunny for a little bit?". Child say this all with the combinations of fastness, quietness, and pauses. And because of a lack or completely unawareness of technical communication skills, a child does not use intonation to show its query as a question. According to researcher, after analysis, the question was good but seems to have multiple questions, and the length is longer than a formal way. Then the mother modifies the child's question and asks, "Alexa, is it going to be rainy all day or sunny all day or . . ." but Alexa couldn't respond this question also because recording breaks then again begin with the noise of the child's laugh, and Alexa has no idea what question was asked and responds "sorry, I'm not sure". Then mother

modifies her question again, and this time she says, “Alexa, I asked if it was going to rain all day.” At this point Alexa understands the question properly and responds with the correct information. Analyzing this example gives us a lot of things about breakdown reasons and their handling. It shows us that the breakdown occurs because of the complex grammatical structure of the sentence, which is unawareness of the child’s ability about the linguistic construction of the sentence. Although this statement can be understood by any human, for any machine or device like Alexa, it’s quite difficult to analyze and interpret the question correctly because of their inability of complex error handling. About mother, she modified the syntax of the child’s question, but Alexa didn’t get it, and in the next attempt, mother modified the query completely in terms of syntax and semantic, which yields the required output. Breakdown occurs from Alexa also, which is known as NR[a1] and explained in table1. And one other reason could also be the code-switching problem; children especially at an early age, are unable to take turns in conversation or use the code-switching methodology. Code switching is the process of shifting from one linguistic code (a language or dialect) to another, depending on the social context or conversational setting [9]. Not only for children but also for home assistant devices, this task is not easy. Involvement of code-switching in training models is very complicated. These algorithms required great effort and time to understand and build the logic completely.

3.2.2 Alexa, You’re Really Stupid

Furthermore, I have explored another research work to illustrate the reasons and types of communication breakdowns and found another paper published in the beginning of 2022, “Alexa, You’re Really Stupid”: A Longitudinal Field Study on Communication Breakdowns Between Family Members and a Voice Assistant.” In this paper, they have conducted a five-week study using Amazon Alexa on ten families, and none of the families have used a home assistant device before. As in the paper described above, they also have children in each family. But unfortunately, they could not get and recover recordings of one family and hence recorded and analysed only nine families’ conversations. The good thing is that they have not only analysed breakdown but also measured user satisfaction level and performed comparisons of children and adults in repair strategies; more precisely, they analysed user behavior. They performed qualitative analysis using regression and, on the basis of hypo-

thetical statements, used some variables for user satisfaction and involvement of emotions like anger and happiness in breakdowns. I found that they discussed children’s interaction with devices can be more prone to errors or breakdowns because of more breaks, grammatical errors, and other unconscious mistakes. Kennedy et al. (2017) [10] described that these characteristics of children directly impact the usage and efficiency of digital devices and make them inefficient. However, the only way to avoid losing existing users or gaining more could be by building more efficient and strongly implemented algorithmic VA that makes sure minimum breakdowns. Actually, this whole research case study was recorded in German and then translated from German to English.

I’ll explain one example from this article that explains communication breakdown in a better way because of its complexity from both sides. This example illustrates how two children tried their best to get some information, but in vain. Although the sentence structure and grammatical context of the question were fine, VA is still unable to answer correctly. Below is the conversation taken directly from the research paper.

Two children, one boy and one girl, ask Alexa which animal is most agile on the Earth. One male child asks the question exactly, “Alexa, what is the most agile animal in the world? ” Surprisingly, Alexa’s response was completely opposite, and it says most venomous animal in the world... and then the girl tries asking a question by making some change and asks, “Alexa, what is the most agile animal in the entire world? ” But again, Alexa’s response is not helpful and says the smartest animal in the world is... Girl tries again with changing sentence structure and question completely and inquires, “Alexa, who is the most agile? ” But this time Alexa’s response was neutral and clearly says I’m sorry... Girl tries one more question, and then boy takes his turn and asks the same question, but every time gets a different response once boy changes sentence also, but Alexa’s response was “The smallest known animal...”, “I like seahorses...”, etc. Once Alexa’s response was neutral and said, “I don’t know, I’m still learning.” Alexa keeps providing the wrong answers even though children seem to apply different breakdown strategies to get the right information. However, as I have explained, before this research was concluded in German and whole recordings were translated into English, that’s why it was difficult to analyze by just relying on English. If we consider the original conversation, we’ll know the problem. It seems to be the problem of speech recognition, as the German words “gelenkigste” mean agile, “giftigste”

Example 3

KM1: "Alexa, what is the most agile animal in the world?"
A: "The sea wasp which is a kind of box jelly is the most venomous animal in the world"
KW1: "Alexa, what is the most agile animal in the entire world?"
A: "The smartest animal in the world is a dolphin. They can recognize emotions, plan own actions in advance and thus, solve complex problems. The capacity of their brains can potentially equal that of the human brain"
KW1: "Alexa, who is the most agile?"
A: "I am sorry, I don't know anything about that"
KW1: "Alexa, is there alien alien something?"
A: "On WELT, there is a documentary 'Aliens, the Message' tomorrow night at 1 a.m. and two episodes on N24 Doku starting from 1:45. By the way, you can ask me questions such as 'What movies are on tonight?' or 'Is there a crime series on today?' More information can be found in your Alexa app"
KM1: "Alexa, what is the most agile animal in the world?"
A: "The longest animal on Earth is a blue whale that is 27.5 m long"
KM1: "Alexa, what is the most agile animal in the world?"
A: "The smallest known animals in the world are the Myxozoa that belong to the phylum of Cnidaria and do not grow bigger than 20 μ m. It is a parasitic lifeform living in the waters. The smallest of the Myxozoa is the Myxabolus shekel with the average size of 8.5 μ m. The smallest vertebrate is the frog Paedophryne amauensis with an average size of approximately 7.7 mm"
KM1: "Alexa, what are the most agile animals in the world?"
A: "Hmm, I don't know, but I am still learning"
KM1: "Alexa, what is the most agile animal in the world?"
A: "I like seahorses"

Figure 2: Image taken from research paper [2] page 10, Example 3 conversation

mean venomous, "längste" mean longest, and "beliebteste" mean beloved. We can see that these words are almost similar to each other and have similar sounds. That's why VA is unable to differentiate this difference and could not predict exactly which words children are saying, providing them wrong answers. In the sentence structure and grammatical context, Animal is a subject, which is not understood by Alexa and yields the wrong output. Children tried their best, and in the end they decided to don't ask again and let it aside. This is the crucial concern for technology usage and the business market. If this remains the same or goes on for a long time, then users will stop using the technology regardless of their techniques and methods to use the device. For example, if a query is repeated again and again, maybe two or more times, then there should be an algorithm that gets the sensitivity of the error and detects something is wrong. The algorithm should be strong enough that instead of giving different responses, as Alexa did in the children's query, it should analyze and catch which word is exactly creating the error that makes the user repeat the same question, and

also instead of giving a neutral response or different answers, it can ask the user about the context of the command, which may reduce the error and can improve user satisfaction and churn. User satisfaction is important and indicates the adaptation and future usage of more similar technology, and if it's not good, it may lead to completely stopping using the technology and making it obsolete. As in the paper, they have also mentioned that according to their hypothetical statement, which seems to be true in regression modeling, the number of unfulfilled failed requests negatively impacts user satisfaction with VA.

3.3 Comparison of Findings

Findings from both research papers identified almost similar issues, including breakdowns caused by Alexa's inability to interpret user input due to semantic or syntactic errors and grammatically incorrect queries given by users. However, the studies also highlighted different user reactions: while some families exhibited patience and tried multiple strategies to repair breakdowns, others quickly became frustrated and abandoned the interaction. These findings suggest that user behavior plays a significant role in how communication breakdowns are handled and whether they are successfully repaired.

3.4 Conclusion

In conclusion, it seems that children make more mistakes in their conversation because of their unawareness of grammatical errors and lack of knowledge because they talk to VA in the same way as they talk to humans and think that it'll understand them properly. On the other hand, children also spent more time to correct their mistakes if breakdown happened as compared to adults. It also shows usage of VA is more sensitive than it seems to be and far more than just a conversation and small tasks like setting alarms, etc. It shows the importance of user satisfaction and how it can lead to dramatic changes in usage of technology, which can make this obsolete if the majority stop using the device because of their errors and more crucial when, despite trying a lot, the user fails to get the right information and rejects the VA at the end. Communication breakdowns either occurred by humans or digital home assistants' need to be captured by devices and that's the part of designing and implementing algorithms. These should be strong enough so that it can understand the context of the

sentence, the difference in similar words and their usage in the sentence, and the accent, although not too much but at least a small variation in the pronunciation. To summarize, breakdowns occur because of semantic and/or syntax error, inability of machine or user or both to code switch, linguistic construction, and unclear instructions. These are not the only reasons for breakdown but are a few out of many possible reasons. We've also analysed the neutral clarification response by Alexa, which could be because of noise during interaction or an unclear question asked by the user.

4 Repair Strategies of Families

In this section, I have explained methodologies adapted by families to overcome communication breakdowns and which methods were most common and useful among families. First of all, I will explain two papers I've discussed before for types of errors; later on, I'll discuss further papers. I have explored 4 different papers, which are explained one by one below to analyze strategies and which methods should be adapted by designers to minimize communication breakdowns and improve overall performance considering common mistakes of users that cause misinterpretation.

4.1 Communication Breakdown between Families and Alexa

According to the explanation provided in "Communication breakdowns between families and Alexa," communication repairs mean efforts to restore common understanding of each other after the occurrence of a breakdown. Mostly the person who has initiated the conversation or speaking needs to paraphrase his wording for better understanding. If we talk about families, that is how they react and solve the error between themselves. This is the case of joint media engagement and discourse scaffolds, as it refers to the guidance and also use of media together. There were six different types of strategies adapted by family members. Asking other family members why they are not getting a proper answer and what they need to ask and how; in short, seeking direction and instructions from other family members. Analogous to this case is "consulting," where family members ask for each other's help to extract some information. If someone extracted something wrong, then articulating his query and adding

more meaningful words in the question to make it clearer and more understandable, is called expansion. Similarly, the opposite of expansion, "contraction," is also adapted if someone has used a too-long sentence that is hard to understand or if there are more irrelevant words, like the case we've discussed before about a 5-year-old child who asks about weather. Another case discussed here for discourse scaffolds is "modeling," which means performing something properly to explain, for example, teaching or showing how to make requests or queries by speaking the query to teach. The last but not least is "Redirection," which shows how to restart the old conversation. For example, if music is playing while Alexa is interrupted for some query, and then again playing that music means redirecting. Table 2 shows the most common methodologies adapted by families to overcome breakdowns according to this paper.

Repair Type	Definition	Example
Prosodic changes	Adjustments to the rhythm or cadence of speech, including pausing and the rate of speech	<i>"Alexa . . what is . . the . . temperature?"</i> [each individual word pronounced slowly and clearly instead of in a conversational manner]- Family C
Overarticulation	Exaggerating sounds, also referred to as hyperarticulation [22]	<i>"Alexa, play 'Make it Rain' by Dack steN"</i> [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]	<i>"Alexa, play little kid music."</i> (Alexa): I couldn't find any little kid songs. <i>"Alexa, play kids songs."</i> - Family G
Increased volume	The speaker raises their voice specifically for the interaction with Alexa	<i>"Alexa stop."</i> <i>"Alexa stop!"</i> [louder] - Family D
Syntactical adjustments	Modifying sentence structure or grammatical elements	<i>"Alexa, thank you, stop it."</i> (alarm continues) <i>"Alexa, stop."</i> - Family H
Repetition	The speaker repeats an utterance used earlier in the same conversation	<i>"Alexa, call [name]."</i> (Alexa): Do you mean [name]? <i>"Volume to eight. Call [name]. Alexa, call [name]"</i> .- Family A

Table 2: Repair Strategies Used by Family Members(Image taken from paper Communication Breakdowns between Families & Alexa from page 6)

4.2 Alexa, you're really stupid

As described above, the repair strategies explained in the second paper are also similar. The most common repair methods were hyperarticulation or prosodic changes, and along these other methodologies were also adapted by different users, like reformulation and repetition.

One other way is to louden their voices, although it doesn't work well. Mostly children adapt these strategies. However, to avoid this breakdown and increase usability, designers need to improve feedback strategies in VAs for better understanding. Keeping all these things in mind, designers need to build home assistant devices that have stronger error handling and repair algorithms instead of focusing more on user satisfaction or demand. Because if a device has better error handling strategy and performance, it'll automatically fulfil user needs. The second most important thing is, as it is also discussed in the first paper, that we can't rely on one side to expect that only technology will improve. Along with technology, we humans also need to learn how we should communicate with machines [11]. It's not like human conversation; we also need to change our conversation style when talking to a machine.

4.3 The 4A's

Now I will draw attention towards methodologies that can be adapted to decrease breakdowns and also describe some other factors that cause miscommunication. The research paper The 4A's Ask, Adapt, Author, and Analyze [12] describes an important term AI literacy that helps designers to understand communication errors in more depth and make them able to design new VAs keeping these factors in mind. For us, the first two terms, Ask and Adapt, are important because they describe how to ask an existing technology and adapt it. The author says it includes many other factors like micro, macro, exo, and mesosystem factors. Microsystem factors imply the learning of people; more specifically, children are affected by their colleagues, classmates, friends, neighbors, and within family learning. The way they learn in this environment has a huge impact on their conversation, which may lead to miscommunication due to their surrounding behavior. Mesosystem factors help us to understand methodologies that are common in communities, like community gatherings, and the impact of one group meeting has a great effect on its adaptation to another group, and direct guidance helps them to improve their learning. At the end, macrosystem factors include social and cultural values and norms. These cultural norms are also applied in ways of conversation that affect the efficiency of VAs either positively or negatively. We also need to consider these social norms that vary from family to family or region to region depending

on their background, and these factors can help us to improve the performance of VAs when making algorithms and training them. This framework also gives guidelines to families to adapt these strategies and understand modern technologies.

4.4 Miscommunication Detection and Recovery in Situated Human-Robot Dialogue

Matthew Marge and Alexander I. Rudnicky have research on Miscommunication Detection and Recovery in Situated Human-Robot Dialogue [13]. This is related to our topic of interest that helps us to find the reasons for breakdowns and how to repair them. It explains communication as a bidirectional aspect that helps us to determine each other's queries and also assists in clarifications to solve misunderstandings. It analyzes features from spoken language inputs the same as VAs. It states that invalid commands given by users are detected by manipulating the existing skills. Language context consists of many components, like recognition of speech and semantics understood from command. These features are presented to the existing speech recognizer vocabulary of the system and perform action by selecting key words, and if there are zero matches from a query, then it is considered invalid or impossible to execute, for which Alexa responds with neutral clarification. For recovery of such instructions from these systems, either the system refashions what it understood according to the context or asks the user to perform their query again. Dialogue systems improve their recovery methods based on previous related interactions. In simple terms, it responds based on its history to how a relevant miscommunication was repaired. To the extent it's ok, but not in all situations, it should also match the context when breakdown happens and adapt different strategies like focusing on syntax and keywords along with the related keywords for better understanding.

4.5 Exploring humor as a repair strategy during communication breakdowns with voice assistants

Mikkel Clausen and his colleagues research on Exploring humor as a repair strategy during communication breakdowns with voice assistants [14]. They experimented that if breakdown

occurs, then VAs use humor as a repair strategy. Humor can play an important role in efficient and successful communication. They have divided humor into four different styles, namely, affiliative humor, self-enhancing humor, aggressive humor, and self-defeating humor, each with its own purpose. As we discussed before in other papers, here we also have some common breakdowns, like NCR, act on misunderstanding, and specific responses. A humor-type questionnaire is used to understand the type of humor and proffered repair choice. This study can help designers apply these strategies, which are more useful and vary person to person and also in context. For example, some people prefer straight-forward responses, while others appreciate the humor to ease their frustration. Another big advantage is we can engage the user with the device to enhance the user experience, and it'll also assist in repairing errors. This also helps in a user-friendly environment even though breakdown occurs. If we consider the context and personal preferences, this will also reduce the risk of miscommunication, hence increasing accuracy.

4.6 Family Negotiation in Joint Media Engagement with Creative Computing

The paper Family Negotiation in Joint Media Engagement with Creative Computing [15] focusses on strategies that families use when breakdown happens, and these methods can be used by designers for better adaptivity of VAs. Joint media engagement means use of a medium like watching TV, playing games to interact together and learn something, or just as a leisure activity. This paper describes a casual creative learning program for families to create a project in an app by name Scratch Jr. during workshops and after getting instructions by facilitators. During this project, families meet with facilitators but not together. Parents and children meet separately and then negotiate together for the project. They can use any kind of story, like a holiday trip. The second phase is to develop the project and share it with other people. Workshops are facilitated by professors and graduates. To analyze how families negotiate and cooperate. It is seen that families negotiate among themselves verbally and physically. Families take turns and take control of the device. Sometimes request to take the device control and at some spots abandon the device, but not forcefully. Children seem to seek help from their parents. Parents help them creatively, like using devices one by one

and asking something related to content, and sometimes parents also encourage children by providing them direct commands that help them to achieve the right information. However, it's also noticed that younger children seem to be not as negotiable as adults and also demand extra devices if someone older is using them for a long time. Three important aspects that are analyzed are the role of parents, evolving strategies, and cultural values and norms. These show that with the passage of time, families improve their interactions on the basis of past experience and results. Thus, strategies change over time. Another important aspect to consider is cultural values and norms and also individual differences, which influence engagement and productivity. And also includes direct involvement and influence of others, especially elders. These strategies can help designers build VAs by considering these cultural norms and adaptive strategies of families, the expectations of families, and what is productive for them. These all will help in developing more efficient VAs for families.

4.7 Comparison of findings

After reviewing all the research papers, I found that few methodologies are common in all for error handling and highlight the importance of adapting communication strategies based on user behavior and needs, such as using simpler language or restating queries (Communication Breakdown between Families & Alexa and Alexa You're Really Stupid). Some of the most common error handling techniques were direct guidance by others or changing the command structure grammatically and semantically. Some other attempts, although they were not most useful, were raising voice and repetition. I found that several studies emphasize the need for improved error handling and recovery mechanisms in voice assistants (Alexa, You're Really Stupid, and Miscommunication Detection and Recovery). One other common factor that influences the users is cultural and contextual considerations. The impact of cultural norms and contextual factors is emphasized across studies (The 4A's and Family Negotiation). Another interesting factor that can improve breakdown is humor. Exploring humor as a repair strategy is unique to the study of humor but aligns with the general aim of improving user experience and engagement. Overall, this essay collectively suggests that effective repair strategies in voice assistants should account for user communication styles, contextual factors, cultural norms, and potentially even humor to improve interaction quality

and reduce miscommunication.

5 Conclusion

In conclusion, this essay describes the type of communication breakdowns that occur between families with digital voice assistant communication and how families repair these breakdowns. I explored different research papers, and under their research, I concluded some reasons for errors and their recovery strategies. As found from the above papers, breakdowns occur due to misinterpretation of the device of incorrect commands given to VAs by users. It is stated that first of all, users need to understand how to interact with machines, and we must keep in mind the conversational difference between humans and agents. Our commands should be more precise and grammatically, linguistically, semantically correct, and interpretable. Machines can't handle all types of errors or mistakes in conversation, thus misinterpret them or respond neutrally. Machines only work on algorithms, yet not so strong algorithms are built that can handle everything easily. This may take time to build strong algorithms that can handle such errors. Overall, it can be concluded that communication errors caused by humans are due to their lack of knowledge of giving commands or rules of grammar. On the other hand, errors caused by machines are due to a lack of understanding of the context and semantics of the query. It is also seen that these are also due to a lack of turn-taking and switching techniques; it can't differentiate rhyming words that are similar to other language words. For that, humans also need to add some keywords, especially in bilingual or multilingual cases. Furthermore, to handle breakdowns, families use either direct involvement of others, which also includes changing the query structure. Some methods that are almost common in all families are louder voice, repeating the same sentence, or repetition with some changes like grammatical or semantics. However, all of them aren't successful. One important aspect that is useful for designers is to keep in mind cultural values and norms, individual differences, and regional characteristics when developing digital voice assistants. Developers need to develop some strong algorithms that handle errors, and instead of giving a neutral response or misinterpreting the query, VAs should ask for the query again by giving some instructions that the query isn't understood by the device, for example, can ask for the word that wasn't matching the corpus. One more thing, that I would like to include, is that adding

some extra rules during the development of VAs, like slang words, and including dialects of regions, may improve the performance and efficiency of VAs. This will also enhance error handling as misinterpretations due to accents and cultural differences will decrease and thus increase user productivity. However, manufacturing something specifically for a region is quite challenging and not easy at all. I believe adapting these all strategies will help digital devices' manufacturers create better VAs that are more prone to errors and will be more demanding, user-friendly, and efficient.

References

1. Beneteau, E., Richards, O., Subramaniam, M. et al. (2020). Communication Breakdowns Between Families and Alexa. Semantic Scholar. Available at: <https://api.semanticscholar.org/CorpusID:140248445>
2. Purington, A., Taft, J., Feaster, J. (2022). Alexa, You're Really Stupid: A Longitudinal Field Study on Communication Breakdowns Between Family Members and a Voice Assistant. ResearchGate. Available at: <https://www.researchgate.net/publication/358181178>
3. Wikipedia Contributors. (collected on August 2024). History of Communication. Wikipedia. Available at: https://en.wikipedia.org/wiki/History_of_communication
4. Fatimayin, F. (2020). What is Communication? School of Education, National Open University of Nigeria.
5. Nakane, I. (2014). Miscommunication and Repair. In: Interpreter-mediated Police Interviews. Palgrave Macmillan, London. Available at: https://doi.org/10.1057/9781137443199_6
6. Suchman, L. A. (1987). Plans and Situated Actions: The Problem of Human-Machine Communication. Cambridge University Press.
7. Luntz, S., et al. (2022). Family Negotiation in Joint Media Engagement with Creative Computing. ACM. Available at: <https://dl.acm.org/doi/10.1145/3544548.3580667>
8. Marge, M., Rudnick, A. I. (2019). Miscommunication Detection and Recovery in Situated Human-Robot Dialogue. ACM. Available at: <https://dl.acm.org/doi/pdf/10.1145/3237189>
9. Springer, L. (2013). Code-Switching in Bilingual Communication*. Britannica. Available at: <https://www.britannica.com/topic/code-switching>
10. Kennedy, J., Lemaignan, S., Montassier, C., Lavalade, P., Irfan, B., Papadopoulos, F. (2017). Child Speech Recognition in Human-Robot Interaction: Evaluations and

- Recommendations. In Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17), 82–90. Available at: <https://doi.org/10.1145/2909824.3020229>
11. Jiang, J., Jeng, W., He, D. (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. Available at: <https://doi.org/10.1145/2484028.2484092>
 12. Druga, S. (2020). The 4As: Ask, Adapt, Author, Analyze - AI Literacy Framework for Families. Available at: <https://api.semanticscholar.org/CorpusID:231626215>
 13. Marge, M., Rudnick, A. I. (2019). Miscommunication Detection and Recovery in Situated Human-Robot Dialogue. ACM. Available at: <https://dl.acm.org/doi/pdf/10.1145/3237189>
 14. Clausen, M., et al. (2023). Exploring Humor as a Repair Strategy During Communication Breakdowns with Voice Assistants. ACM. Available at: <https://doi.org/10.1145/3571884.3597127>
 15. Luntz, S., et al. (2022). Family Negotiation in Joint Media Engagement with Creative Computing. ACM. Available at: <https://doi.org/10.1145/3544548.3580667>

Abbreviations

- I VAs – Voice assistants
- II NCR – Neutral Clarification Response
- III CMB – Communication Breakdowns
- IV JME – Joint Media Engagement
- V HCI – Human Computer Interaction

List of Tables

1. Signals from Alexa that a communication breakdown has occurred, in Communication Breakdowns between Families & Alexa page No. 6, Table 4
2. Repair Methodologies adapted by Family Members in Communication Breakdowns between Families & Alexa from page No.6 Table 2

List of Figures

1	Image taken from research paper [1] page 9, Family D conversation	6
2	Image taken from research paper [2] page 10, Example 3 conversation	9