What do you mean? A Review on Recovery Strategies to Overcome Conversational Breakdowns of Conversational Agents

Completed Research Paper

Dennis Benner
University of Kassel, Germany
dennis.benner@uni-kassel.de

Edona Elshan
University of St. Gallen, Switzerland
edona.elshan@unisg.ch

Sofia Schöbel
University of Osnabrück, Germany
sofia.schoebel@uni-osnabrueck.de

Andreas Janson
University of St. Gallen, Switzerland
andreas.janson@unisg.ch

Abstract

Since the emergence of conversational agents, this technology has seen continuous development and research. Today, advanced conversational agents are virtually omnipresent in our everyday lives. Albeit the numerous improvements in their conversational capabilities, breakdowns are still a persistent issue. Such breakdowns can result in a very unpleasant experience for users and impair the future success of conversational agents. This issue has been acknowledged by many researchers recently. However, the research on strategies to overcome conversational breakdowns is still inconclusive, and further research is needed. Therefore, we conduct a systematic literature analysis to derive conceptual conversational breakdown recovery strategies from literature and highlight future research avenues to address potential gaps. Thus, we contribute to theory of human-agent interaction by deriving and assessing recovery strategies and suggesting leads for novel recovery strategies.

Keywords: conversational agents, conversational breakdowns, recovery strategies

Introduction

Conversational agents (CA) have been around for many years by today; however, the real hype about CAs, some driven by artificial intelligence (AI), only became largely popular in the 2010s of this century (Brandtzaeg and Folstad 2018). Nowadays, conversational technology is virtually omnipresent in different artifacts of our everyday lives, such as text-based support or educational chatbots and voice assistants like Amazon Alexa or the Google Assistant on our phones and in our smart homes. CAs are also gaining increased exposure and attention in practice, with CAs being included on many commercial websites and in digital services. These examples, however, cover only a small scope of today's usage of CAs. Today, CAs have been deployed in a variety of domains, such as traveling, shopping, recommendation system, entertainment, customer service, information search processes, and so on. In general, two main types of CAs can be distinguished: Social-oriented CAs and task-oriented CAs. Social-oriented CAs focus on social interactions and managing emotions and other informational inquiries, whilst task-oriented CAs support the user in specific tasks (Xiao et al. 2019). Especially the latter, task-oriented CAs, turned out to be extremely helpful in performing customer support tasks as they bear the potential to reduce the amount of human labor needed considerably.

Besides the omnipresence of CAs, interacting with them oftentimes results in questions such as "*What do you mean*?" leading to an abrupt ending of a conversation. If the CA is not able to handle the users' conversation smoothly, breakdowns in the conversation might lead to the user abandoning the service (Ashktorab et al. 2019). Researchers and practitioners have dedicated themselves to improve the interaction between humans and CAs, for instance, by developing better natural language processing techniques (Meredith 2017). Despite these improvements in the field of CAs during the past years, there are still many unresolved issues (Ashktorab et al. 2019). For example, we still lack the means to adequately handle natural language conversations between users and CAs. In addition to natural language or linguistic errors in general, bad experiences due to poor interactions and conversational errors (e.g., CA responds with an unexpected action, ignores an inquiry, tiresome and boring dialogue), the user can become demotivated quickly (Brandtzaeg and Folstad 2018). Such a quick demotivation can then equally quickly lead to a complete conversational breakdown and consequently a complete failure of interaction (Brandtzaeg and Folstad 2018). Therefore, keeping the user motivated and engaged becomes an important aspect of designing conversations and interactions with agents. All these possible errors can eventually lead to total rejection and loss of trust in CAs (Engelhardt et al. 2017). Nevertheless, CAs are able to maintain a themed discussion close to 85 percent of the time (Radziwill and Benton 2017). As encounters within social or commercial contexts rely on the user being engaged and having a pleasant experience, the remaining 15 percent of the conversations are vital. Against the backdrop that breakdowns might lead to annoyance (Chakrabarti and Luger 2015), awkwardness, or confusion (Bickmore et al. 2018). All these negative factors may then result in a loss of trust in the CAs capabilities and users breaking up the conversation with a badly designed or faulty CA. As service encounters rely on CAs, the interaction with the CA influences a customer's buying decision (Shankar et al. 2003) and has various effects such as on the word of mouth and intention to buy a product repeatedly (Bitner et al. 2000).

Therefore, it is important that we define a working set of strategies to effectively handle conversational breakdowns when they occur, as overcoming conversational breakdowns is critical to the success of CAs. Prior research sought to uncover when and why communication errors and failures like mis- and nonunderstanding occur. For instance, in their study, Li et al. (2019) showed that conversation breakdowns appear more often when the users were requesting information than when they were providing information. Skarbez et al. (2011) revealed that the CAs' inability to recognize their role as speakers or listeners contributes significantly to conversation breakdowns. Furtherf, previous studies show that the conversations serve their purpose with a very straightforward approach. Nevertheless, they do not offer any surprises that make a conversation exciting in the first place (Black 2018). The result is CAs that fail and respond with standard statements such as "*I'm afraid I didn't understand* that". It was recognized early on that in human-computer interaction, and human characteristics might also be also attributed to the computer and thus to CAs as well. To improve communication between humans and CAs, there is an attempt to add human-like social attributes to the conversations (e.g., small talk) to increase the likability of agents. (Blut et al. 2021; Clark et al. 2019).Other studies also emphasize that trying to make the CAs more and more human is not even necessary (Radziwill and Benton 2017). Although initial literature reviews and meta-studies (i.e., Janssen et al. 2020; Xiao et al. 2019) emerged during the past years, the research is still scattered across different streams of research. Overall, the scientific and practical knowledge about how to overcome conversation remains segregated. The goal and contribution of our research are twofold. First, we will provide an overview of current methods of adaption and recovery strategies based on a systematic literature analysis with a wide and explorative scope. Second, we will provide future directions for research on this topic and potential options that could prove as a valuable addition to improving recovery strategies for CAs. Accordingly, we will answer the following research question (RQ) in our paper:

RQ: What is the state-of-the-art concerning strategies for recovery of conversation breakdown for interactions with conversational agents?

To answer this research question, we conceptually derive a set of recovery strategies for conversational breakdowns. In doing so, we contribute to theory by providing an overview of the status quo on conversational breakdown recovery literature and extend this knowledge by adding a set of six recovery strategies with unique characteristics. Moreover, we also hope to discover promising research avenues for future contributions and support practitioners in overcoming breakdowns of CAs. The remainder of this paper is structured as follows: after motivating our research idea, we provide an overview of conversational agents and elaborate on our understanding of miscommunication as well as conversation breakdowns.

Next, we describe our research approach and present the results of our literature research before we discuss them. We conclude with a short outlook and our contributions to research and practice.

Theoretical Background

Conversational Agents as Conversational Actors

Generally, CAs can be classified as either physical or virtual autonomous technological entities that can behave reactively and proactively (Holz et al. 2009) by accepting natural language as input and by generating natural language as output to engage in social conversations with its users (Griol et al. 2013; Keyser et al. 2019). Various terms can be found in literature for conversational agents (CAs) such as chatbots, voice assistants, voice bots, or smart personal assistants. Lieberman (1997) defines CAs as software programs that act as a facilitator or an assistant rather than a tool. Also, because CAs are computer programs that interact with humans using natural languages, and their goal is to simulate human conversation (Bittner et al. 2019). Due to their conversational, human-like character, CAs bear great potential to access content and services in a more intimate and personal manner than usual, non-conversational self-service technologies (Sheehan et al. 2020). They also hold the promise of providing a fast, easy, always available, and cost-effective solution to support users (e.g., Holz et al. 2009). Moreover, unlike traditional self-service technology (e.g., web interfaces), users can better relate to CAs because of their aforementioned human-like characteristics creating a more personal and likable service delivery. Consequently, it is to no surprise that CAs are now employed in a wide range of application areas such as health (e.g., Laumer et al. 2019), education (e.g., Winkler et al. 2020), and customer service (e.g., Qiu and Benbasat 2009). Bearing that in mind, the industry expects the users' adoption to be relatively high and the usage of self-service CAs to increase (Nordheim et al. 2019). However, the industry's expectations do not necessarily meet the reality, as the opposite can be true. Particularly error-prone CAs that cannot sustain a solid conversation can harm the intent of the users to adopt these CAs, leading to significant issues for the self-service industry (Sheehan et al. 2020). Due to the aim of providing a flawless user experience, the development of CAs is a challenging task for their designers. Within the development process, they need to account for the user's preferences regarding the assistant's personality and talking style and prepare a corresponding script for the interaction with the CA. In addition to that, they have to anticipate the user's actions and derive appropriate reactions from the CA.

Conversational User Experience Design

The way we interact and communicate with systems has changed continuously over the years (Folstad and Brandtzaeg 2017). Interfaces that allow us to exchange information with computers evolved from command-based command-line interfaces to graphical user interfaces to modern augmented reality and conversational user interfaces (Wintersberger et al. 2020). In the context of CAs, interaction takes place via a conversational user interface, which enables text or voice input and is kept relatively minimalistic compared to graphical user interfaces (McTear 2017). Due to the most simplistic conversational user interfaces, the challenge in developing CAs is not in designing the interface but in designing an effective and stimulating user experience (Sutcliffe 2009). Accordingly, graphical elements are less relevant than the conversation flows and language capabilities of the CAs. In CAs, human-computer interaction does not take place via gestures, such as clicking or swiping, but by stringing together text modules (Brandtzaeg and Følstad 2017). Moore and Arar (2019) describe conversational user experience design in this sense as "[...] *modeling the patterns of human conversation, either through manual design or machine learning*" (Moore and Arar 2019, p. 3). To do this, CA developers need to build a deep understanding of human language and understand what typical elements and building blocks occur in classic conversations. For this reason, conversational user experience design is shaped in diverse aspects by disciplines such as sociology or psychology (Moore and Arar 2019).

When it comes to conversation breakdowns, two different kinds of conversation breakdowns can be classified: misunderstandings and non-understandings (Bohus and Rudnicky 2005). Whilst in a misunderstanding, the CA obtains an incorrect interpretation of the human's input, non-understanding leads to failure of obtaining any kind of interpretation of human's input, although both types of failure can and will eventually lead to total CA failure (Bohus and Rudnicky 2005). Further, by means of conversational analyses, three typical patterns in human speech behavior could be identified. These are described by Moore and Arar (2019) as Recipient Design, Minimization, and Repair. These fundamental principles of conversational user experience design should be considered in any case of CA design, as their use can make human-computer interaction more natural. The adapted principles are each briefly explained below.

Recipient Design: In a regular conversation, both parties adjust their dialogue depending on the conversation partner; different topics, words, and levels of details are chosen, e.g., user-friendly language with the help of cues. Because this mode of conversation assumes prior knowledge of the parties, misunderstandings can occur and consequently lead to conversational breakdowns (Moore and Arar 2019). Thus, recovery strategies should include adaptable conversation flows to create multiple dialogue paths to prevent or counteract a breakdown. For instance, Fast et al. (2018) introduce a CA that asks follow-up questions and presents plausible conversation pathways when it comes to mis- or non-understandings.

Minimization: In most cases, conversations between people proceed efficiently. Conversations are not unnecessarily prolonged, taking up the time of all parties and feeling tiresome. For the conversation design of CAs, it can be deduced from this that the statements in interaction with CAs should basically be formulated as briefly and concisely as necessary (Clark et al. 2019). Thereby, care must be taken to ensure that as many users as possible understand the statements; thus it should not be overly abstracted. So, minimizing required dialogue while maximizing its effectiveness, ensures that mis- and non-understandings are minimized as well. In the context of recovery strategies, this implies that these strategies must also minimize the necessary dialogue.

Repair: It is natural that in conversations, some statements or intentions are not understood at all, or only partially. In such cases, humans can correct difficulties in understanding by formulating misunderstood statements more simply or by repeating them altogether. In conversation analysis, this principle is called repair. The behavioral mechanism of identifying and correcting problems of mutual intelligibility can be initiated by both the person speaking and the person receiving. For conversation design, appropriate improvement loops must be considered, which allow both the CA and the user to fix misunderstood statements or ambiguities (Moore and Arar 2019; Schuetzler et al. 2020). Recovering a conversation may involve the CA asking the humans to repeat themselves, to refer to prior information, and/or to clarify what they said (Bickmore et al. 2018). Further, prior research shows that users are less frustrated if systems such as CAs apologize for errors (Bulyko et al. 2005).

Breakdowns and Recoveries in Human-Agent Communication

Since interactions during communication may not be understood or misunderstood, communicative interactions can easily fail. This is particularly important for the development and design of CAs, as people who use them usually expect them to work just like other technological artifacts they know. Thus, the importance for recovery from conversational breakdown and strategies to attempt to repair the faulty conversation. In general, repair is broadly defined as the replacement of errors with corrections (Schegloff et al. 1977). In this context, recovery attempts may be initiated by either the CA or the user. On the one hand, the user may simply rephrase or reformulate the request to the CA that did not properly understand the users' intentions, hence attempting to recover. On the other hand, the CA may have a mechanism in place that allows the agent to detect a potential conversational breakdown and thus act according to attempt a recovery. For example, the conversational agent may ask the user to repeat or rephrase the question.

A very common reason for such a conversational breakdown is the failure of the underlying natural language processing and interpretation by the conversational agent (Myers et al. 2018). While the research on the topic of natural language processing in this regard was and still is a significant research interest, research on alternative mechanisms to recover from breakdowns seems to play a minor role. Not surprisingly, conversational breakdowns mark a long-lasting concern and limitation of conversational agents and their effectiveness in human-agent communication (Luger and Sellen 2016).

Research Approach

To answer our research question, we conducted a systematic literature analysis (SLA). The analysis is based on established literature review methods as introduced by Cooper (1998), Fettke (2006), and Vom Brocke et al. (2015). Below, a simplified version of our structured literature analysis process is shown in Figure 1. The process is adapted from Fettke (2006). At the start of our SLA, we define the characteristics of our review and classify the SLA according to Cooper (1998). The goal of our SLA is the identification of central issues in the status of the art and the generalization of existing solutions to the common denominators, which according to Cooper (1998), fits the integration goal category. Accordingly, the focus of our SLA is set on the theories, practices, and applications of the solutions as well as the outcomes thereof. The classification is shown in Table 1.

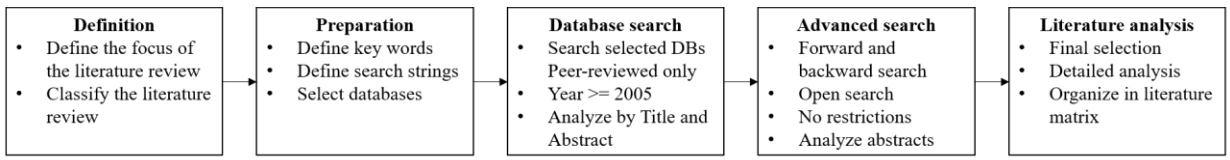

Figure 1. Process of the SLA (adapted from Cooper 1998 and Fettke 2006)

As our aim is to provide a generalizable overview of recovery strategies for CAs, both voice-based and text-based, thereby our coverage is exhaustive. Further, the organization of our findings is strictly conceptual as we also employ the literature review methodology as suggested by Watson and Webster (2020) and present a structured concept matrix in the results section. Our target audience hereby is specialized scholars in the field of human-computer interaction (HCI) that focus on designing CAs.

Characteristic	**Category**		
Goal	Integration	Criticism	Central topics
Scope	Representative	Selective	Explorative (complete)
Focus	Results & outcomes	Methods & designs	Theories
Structure	Historical	Thematical	Methodical

Table 1: Taxonomy of the literature review (adapted from Cooper 1988)

As a second step of our SLA, we prepare the database search process. Therefore, we define desirable keywords and construct search strings that we will use for the database search, as well as selecting relevant databases. Because this topic is rooted in computer linguistics, we chose the ACM Digital Library and the IEEEXplore Digital Library as databases to acknowledge for this in our literature review. We also include the AIS Electronic Library as many outlets highly relevant to the research field of HCI, and generally, high-quality outlets are hosted by the AIS. The keywords and consequently the constructed search string is shown below. The search string was adapted and simplified to the characteristics of each database, with wildcards being used where possible.

("conversational agent" OR "chat bot" OR "chatbot" OR "dialogue system" OR "smart personal assistant" OR "smart assistant" OR "intelligent agent" OR "intelligent assistant" OR "spoken dialog system" OR "conversational spoken language interface") AND ("error recovery strategy" OR "error recovery strategies" OR "restart strategy" OR "restart strategies" OR "error detection")

Following the preparation phase, we conduct our database search process with the defined search string and analyze the title and abstract of all research papers. The only two restrictions we impose on the research process are (1) the literature must be peer-reviewed and (2) the literature must not be older than 2005. As a result of our primary search process, we identified 136 articles that we found to be potentially relevant. We examined those articles by full text in detail. In a next step, we performed a forward and backward search to capture articles not covered through the database search. Through screening the references and applying forward searches using Google Scholar, three papers were added to our literature analysis. In the end, we kept only those articles that cover the topic of strategies of conversational recovery, repair or any other attempt to salvage a conversation. Here we excluded all articles that covered only technical implementations of strategies or mechanisms, such as highly technical articles that are more concerned with the algorithmic implementation rather than conversational design or engineering. As an end result of our search process, we kept 31 articles that we found to be relevant for our research. In the fourth and last phase, we conducted the literature analysis by following an abductive approach. The coding process was iterative and required multiple rounds of different researchers until we were able to organize the literature in a conceptual matrix. In the course of the literature analysis, we captured the recovery strategy and tried to code them accordingly as a reaction to misunderstanding or non-understandings. Furthermore, we considered non-verbal as well as verbal recovery strategies in our coding. Additionally, we captured contextual variables such as the modality, the application domain, and task of the CA. Subsequently, we discussed how to combine recovery strategies across studies, which lead to our seven breakdown strategies: confirmation, information, disclosure, solve, ask, and not specified.

Results

We organize our findings in two sections. The first section presents an overview and the descriptive metrics based on the meta-data of the underlying literature and technical characteristics of CAs. Thus, we start by presenting our SLA findings and the status quo of the research on conversational breakdowns. The literature we analyzed is organized in a literature matrix (see Table 2) according to Watson and Webster (2020). Accordingly, we organize the literature in two meta-categories: 1. technical properties that describe the technicalities of the CAs presented in the papers and 2. the breakdown recovery strategies that we derived from the literature. Additionally, we describe the nature of the research (i.e., theoretical vs. practical research). In the second part of our results, we present the derived conceptual conversational breakdown strategies in detail (see Figure 4). The second section then examines the recovery strategies for conversation breakdowns used in CA research in detail.

Status Quo about Recovery Strategies

The youngest paper included in the analysis is from 2020, and the oldest paper from 2004, from before CAs became widely popular. The work of Sagawa et al. (Sagawa et al. 2004) was found during the forward and backward search as our literature search process focused on papers from at least 2005. Nevertheless, the majority of papers have been published within the last five years, which highlights the continuous development and expansion of the research field concerning CAs. Additionally, the research on the topic of conversational breakdowns and recovery strategies has gained significant momentum since 2017 as about half the found literature was from at least 2017. Since 2017 this research interest has grown extensively with a steep and linear surge in papers published. This research development and the distribution of the relevant literature we found can be seen in Figure 2 (number of articles by year of publication).

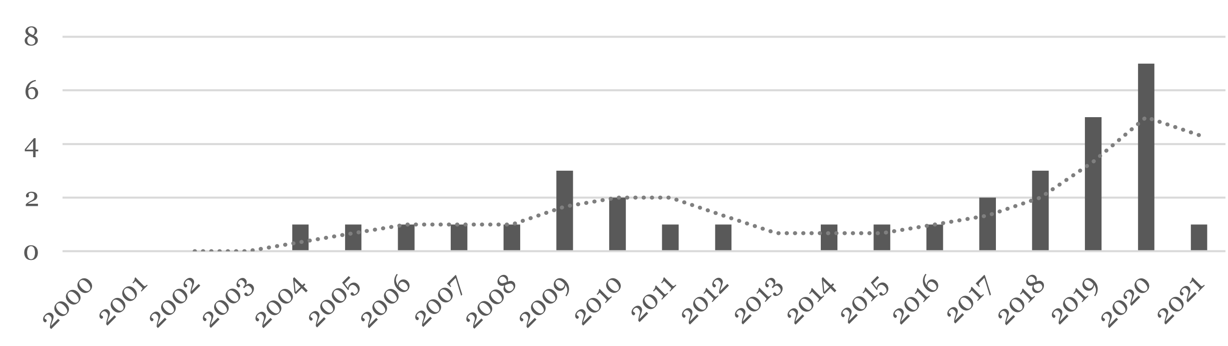

Figure 2. Distribution of Relevant Literature according to Year

However, we also found relevant literature from before the recent spark in research on conversational breakdowns. Older research is evenly distributed from the year 2004 up to 2011 with a small but steady interest in the topic. The years of 2012 and 2013 seem to be an exception in our SLA as we did not identify relevant literature from these years. Further, we could not find a single relevant paper from before 2004, which is not surprising as the topic did not gain a notable amount of attention before 2005. This may also be explained by the early expansion and awareness around CAs in the late 2000s up to 2010.

	Technical Properties						Recovery Strategies						Research	
	Communication channel		Understanding error type		Recovery strategy type		Confirmation	Information	Disclosure	Social	Solve	Ask	Theoretical or conceptual	Empirical
Author/s (year)	S	T	N	M	V	X								
Ashktorab et al. (2019)		X	X	X	X		X				X	X		X
Bohus and Rudnicky (2005)	X		X	X	X		X	X				X		X
Dzikovska et al. (2009)		X	X		X			X				X		X
Dzikovska et al. (2010)		X	X		X			X		X				X
Engelhardt et al. (2017)	X		X	X	X	X		X		X				X
Frampton and Lemon (2008)	X		X			X					X	X		X
Gnewuch et al. (2017)		X		X	X		X	X					X	
Gunkel (2016)	X		X	X	X					X			X	
Lee et al. (2018)	X		X	X	X			X		X				X
Lee et al. (2020)		X	X	X		X						X	X	
Li et al. (2019)		X	X	X	X							X		X
Li et al. (2020)		X	X	X	X	X					X			X
Kocaballi et al. (2020)		X	X	X	X	X		X				X	X	
Komatani and Okuno (2010)	X		X		X			X						X
Komatani and Rudnicky (2009)	X		X	X		X					X			X
Komatsu and Sasayama (2019)	X				X					X				X
Kontogiorgos et al. (2020a)	X		X		X		X			X		X		X
Kontogiorgos et al. (2020b)	X				X	X				X				X
Lee et al. (2007)	X		X	X	X							X		X
Lee et al. (2014)	X		X	X		X		X						X
Litman et al. (2006)	X		X	X	X		X			X		X		X
Marge and Rudnicky (2019)	X			X	X		X				X			X
Mozafari et al. (2020)		X			X				X					X
Mozafari et al. (2021)		X			X				X	X			X	
Porcheron et al. (2018)	X		X		X							X		X
Rieser and Lemon (2011)	X		X			X	X					X		X
Sagawa et al. (2004)	X			X								X		X
Sheehan et al. (2020)	X			X	X		X			X				X
Stoyanchev and Stent (2009)	X			X	X		X							X
Stoyanchev et al. (2012)	X			X	X							X		X
Stoyanchev and Johnston (2015)	X			X	X							X		X
Takayama et al. (2019)		X	X	X		X					X			X
Woodward et al. (2018)	X		X		X			X	X					X
Sum (*n* = *33*)	22	11	22	21	26	10	9	10	3	10	6	15	5	28

S = speech or voice; T = text-based; N = non-understanding; M = misunderstanding; V = verbal; X = non-verbal

Table 2. Analysis Matrix of Conversational Breakdown Recovery Literature.

Additionally, this development also shows that researchers are becoming more aware of the issues concerning conversational breakdowns. This is also reflected by the fact that most papers are from conference proceedings, which gives testament to the still relatively low maturity of this research field, especially regarding conversational breakdowns and recovery strategies. Historically speaking, the first contributions are rather explorative, incorporating a multitude of investigated conversation breakdowns, while recent papers are more specific concerning the occurred conversation breakdown and recovery strategy applied. Further, the examined contributions included studies from various application contexts and data sources. Moreover, early literature does not specifically tie the research to conversational agents, instead of the terms of "(speech/spoken) dialogue systems" (e.g., Lee et al. 2007; Sagawa et al. 2004) and is notably focused on the technical or computational aspects of handling conversation. Regarding the type of research, the vast majority of research is focused on empirical or practical research like field studies (28), laboratory experimentation of other data-based efforts, instead of theory-based research (5) as literature analysis, theory crafting, or in-depth discussions.

Concerning the technical properties, we describe the communication channel, the error type of the understanding error, and the recovery strategy type of the CA. For the communication channel, we identified voice or speech (S) and text-based (T) CAs. Since we also include robots, these instances are classified as voice or speech, as all included robots communicate with humans via mostly voice or speech and, in some rare occasions, also text channels. The understanding error type refers to the previously described possible errors in the communication between humans and CAs, namely non-understanding (N) and misunderstanding (M). The recovery strategy type precedes the breakdown recovery strategies and refers back to the communication channel. Recoveries can be attempted verbally and non-verbally, where verbal recoveries refer to both speech or voice and text-based communication. Non-verbal recovery, however, does not directly communicate with the user to execute a recovery strategy. Instead, CA internal measures can be employed like recalculating scores or changing internal utterances and conversational templates that are used to determine the CAs actions. Most of the literature we surveyed uses CAs that use voice or speech as a communication channel (23), only a minority of six CAs is text-based. The majority of CAs also use verbal recovery strategies (24), whereas only a minority of CAs use non-verbal strategies (10). However, the regarded understanding error types are much more equally distributed, with 22 focusing on non-understanding and 20 on misunderstanding. This may imply that while research is, in general, focused on speech- or voice-based CAs that use verbal recovery, researchers are split on the error type. Some researchers focus on both types (4) as seemingly both are crucial to the success CAs.

Deriving Strategies for Conversational Breakdowns

As for the breakdown recovery strategies, we identified six distinct categories with distinct characteristics: (1) confirmation, (2) information, (3) disclosure, (4) social, (5) solve and (6) ask. The distribution of the six categories is relatively even, with disclosure being underrepresented and ask an overrepresented exception in literature. As we derived these recovery strategies from literature, we found them to be differently distributed and implemented by the authors. For one, information and social recovery strategies with ten and with ask with 15 findings are the most prominent strategies from literature, whereas both disclosure (one finding) and solve (one finding) being less prominent exceptions. We describe these six categories along with their common mechanics and characteristics. These categories can be seen in Figure 3.

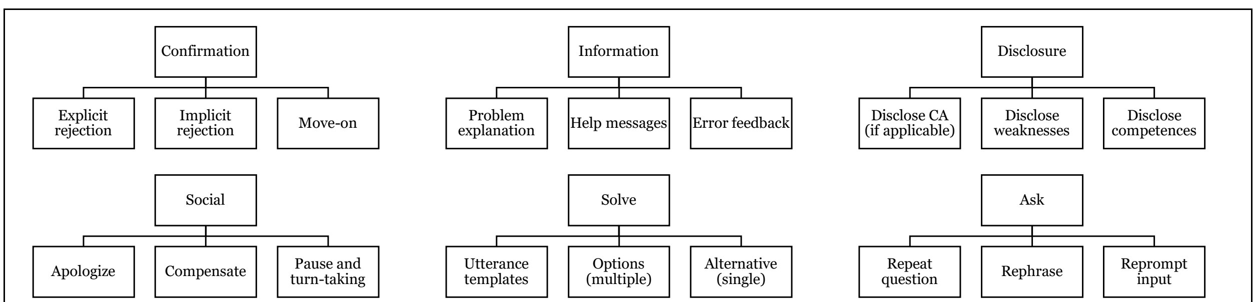

Figure 3. Concept Map of Conversational Breakdown Recovery Strategies

Each category may consist of two or more subcategories that may be expanded in the future. Here, we found three subcategories for each major category of conversational breakdown recovery strategy. We now present the six major categories in detail as following.

The **Confirmation** strategy does not attempt to truly recover a failed conversation. Instead, the confirmation strategy attempts to either reject the failure or simply "move on". In this regard, rejection can be either implicit or explicit. On the one hand, explicit rejection directly translates into the CA admitting its failure. For example, the CA may respond with "I do not know" or "I do not understand," and thus, hoping for the user to attempt a recovery. Implicit rejection may be expressed as merely ignoring the failure and hoping for the user to detect and recover from a breakdown. The move-on subcategory expresses itself basically as "ignore and continue", where the CA ignores the failure and tries to continue in a pre-programmed path. For example, Woodward et al. (2018) state that CAs should explicitly admit and confirm failure when they are not able to provide a reasonable response to the user. In fact, Woodward et al. (2018) emphasize that CAs should simply reply with "I don't know" when they really are not able to provide an otherwise satisfying response.

The **Information** strategy is more advanced in comparison to simple confirmation. The CA will attempt to explain the situation, provide potentially helpful messages or feedback on the error. For example, if the CA is programmed to only accept a certain set of characters or a certain language type like strictly formal language, the CA may fail to understand informal language (e.g., slang) and thus explain to the user that the CA requires a specific input to be useful. For example, Dzikovska et al. (2010) use two different policies. In the first, a misunderstanding was assumed in case of a user was always given the correct answer to the question in response to the utterance. In the second policy, the user was given a help message with a hint to the correct answer in case of an incorrect answer.

Disclosure is similar to the information strategy in that it seeks to educate the user on how to solve a potential breakdown. However, there is a significant difference. Disclosure does not specifically and directly address the error or breakdown. Instead, the CA may disclose itself as a CA or computer artifact and exposing its weaknesses and competencies. For instance, a CA may be perceived as a real human being and set unrealistic expectations; thus the agent may disclose itself to the user and, in doing so, set realistic expectations as the user will know that the counterpart does not possess human intelligence (Grimes et al. 2021). Further, by disclosing weaknesses and competencies, the user will be informed of the skills and potential shortcomings of the CA and, thus, further refining the users' expectations and understanding of the CA. This may result in a higher acceptance as the user will exactly know what the CA is good for and what not. For instance, Mozafari et al. (2021) studied the impact of disclosing to users that the service encounter is provided by a CA. No additional studies in our dataset were found that looked more closely at this category of recovery strategy. Thus, this could be a potential direction for future research.

The **Social** recovery strategy addresses the human aspects of human-agent conversation by introducing typically human behavior into the dialogue. One expression of this is the CA apologizing for the breakdown or errors and, in doing so, trying to appeal to the users' empathy and understanding similar to that which is shown in human-human conversations (e.g., Engelhardt et al. 2017). Additionally, the CA may try to compensate the user. In this case, the compensation may be emotional (e.g., further apology or inclusion of social cues) or non-emotional by, for example, offering other incentives to continue the conversation and attempt to recover. As a third expression of the social strategy, the CA may partake in pausing and turn-taking, which is similar to human-human conversation where parties will sense when to converse and when not to (e.g., Gunkel 2016). Currently, there are many efforts in research to make CAs more human through the application and implementation of social design elements (Zierau et al. 2020) or gamifying design elements (Benner et al. 2021). This certainly includes the reaction to conversation breakdowns. Therefore, the category social seems to be vital within the breakdown recovery strategy. What is striking, however, is that this category is hardly ever used on its own.

Solve as a strategy is goal-oriented in the sense that the CA tries to actively solve the breakdown by providing a solid solution. One example of this strategy is to provide an alternative that is most likely to what the user wants based on the conversation so far. This recovery strategy is similar to "taking an educated guess" in human-human conversation. Similar to this approach, the CA may also provide the user with a list of options that potentially solve the problem. Moreover, the CA may also offer pre-programmed utterance templates to the user and encourage the user to use them to attempt recovery. For example, when the user uses some type of language or expression the CA cannot process, the CA may offer a set of utterances that the CA will certainly understand. Such a recovery strategy can be observed within the study of Marge and Rudnicky (2019). Lee et al. (2007) employ so-called "utterance templates" as breakdown recovery strategies. These templates include specific instructions and utterances for the user to execute to recover from a conversational breakdown. The crucial distinction from information is that information does not provide specific measures to solve a breakdown.

The **Ask** strategy is very simple and straightforward. Using this recovery strategy, the CA shifts the burden of recovery to the user by employing three different techniques. Firstly, the CA may simply repeat the question and thus allow the user to make another request (e.g., Li et al. 2019). Secondly, the CA may ask the user to rephrase the request or sentence (e.g., Lee et al. 2007). Thirdly, the CA may (re)prompt the user for input, for example, if the first try was unsuccessful or the CA requires more information to provide a satisfactory answer (e.g., Stoyanchev et al. 2012). As users are prone to use ambiguous concepts and expressions when interacting with CAs, it is only natural that the CA shall ask for clarification (Li et al. 2019). Additionally, sometimes words can be very closely related and even consist of the same letters, like for example, the words "Tokyo" and "Kyoto" which CAs may confuse or not recognize properly since their makeup of letters and their lengths is the same, just the ordering is not (Sagawa et al. 2004).

These six recovery strategies for conversational breakdowns in human-agent conversation used individually or in combination. For instance, Litman et al. (2006) used a combination of the categories ask and social. In this case, the CA replied, "Sorry I can't understand you, can you please repeat your utterance?", when it was facing a breakdown in the conversation. Another combined strategy can be found in the study of Engelhardt et al. (2017). After the CA has apologized to the user for non-understanding him/her, the CA provided the user with potential problem-solving options. Although we have divided our paper and the recovery strategies mentioned therein into the aforementioned six categories, it is certainly feasible and likely that additional recovery strategies may emerge in the future. In the following discussion, we present three examples of potential future research avenues to further develop conversational breakdown strategies and expand our presented conceptual recovery strategies.

Discussion and Contributions

This study aims to advance the state-of-the-art conversational breakdown recovery strategies for human-agent interaction. As a result of analyzing and grouping those strategies in a concept matrix, we provided a comprehensive view of them. We can observe the growing significance of these strategies since research is showing continuous growth during the last few years, as well as the importance to gain a deeper understanding of the outcomes of their implementation in a wide range of contexts as the research seems to be in a rather early stage. In this section, we discuss and provide future pathways for research, beginning with the state-of-the-art strategies identified in this literature review.

Circling back to the initially presented three criteria by Moore and Arar (2019), which essentially refer to the CAs ability to adapt to changing conversation flows (1), efficient conversation minimizing the amount of dialog (2) and using effective behavioral mechanisms to identify and correct (i.e., prevent or recover) conversational problems, we presented our recovery strategies. These strategies adhere to these three criteria but differ in their potential use for varying scenarios or configurations of CAs. Researchers and practitioners need to be able to assess the use of these strategies. Thus, we want to present our assessment of conversational breakdown recovery strategies and their potential use for different scenarios and configurations of human-agent interaction. So, we compiled an overview of our assessments for potential future additions to recovery strategies (see Table 3), which we deem as valuable.

The first criteria we base our assessment on is the action type. A CA can either actively (e.g., admit failure or apologize) or passively (e.g., using cues or hints) try to recover from a breakdown. For an active recovery, we assess that all strategies other than persuasion may be applied because direct and active persuasion may be perceived as disturbing or disrespectful. We assess that confirmation, solve, ask, and handover are not possible for passive recovery, as all these strategies require active intervention by the CA. Information and disclosure may be useful if applied in a subtle manner. For passive recovery, persuasion may prove useful as persuasion itself should be applied in a subtle way to not disturb the user. For the channel, we assess that all strategies are useful for text-based communication with some limitations on the social strategy as speech can transmit potentially rich information during interaction (e.g., tonality). For voice communication, we find that all strategies, but disclosure and handover are suitable. Disclosure may be used to inform about strengths and weaknesses but not to disclose the CA as the user will most likely already know (e.g., Alexa). Handover is a difficult topic in this regard, as users may experience social anxiety and prefer easier text-based communication with real humans. Concerning the error type, we assess that only confirmation, solve, and ask may be useful strategies due to the difficult nature of misunderstandings in communication, which is often true to human-human conversation. For example, providing alternatives (e.g., earlier example "Tokyo" vs. "Kyoto") seems reliable, whereas disclosure or handover may not be, as the CA may not be able to comprehend the issues and its own mistake leading to the misunderstanding. For non-understandings, we assess basically all strategies to be useful in one way or another.

Based on the results of our study, we can give implications for future research projects, namely: about the handover of a CA, the supported user group, and motivating design components of CAs. We explain our thoughts on the potentials for these strategies in the following.

Strategy	Action type		Channel		Error type		Addition
	Active	Passive	Text	Voice	Mis-understand	Non-understand	
Confirmation	+	–	+	+	+	+	No
Information	+	o	+	+	o	+	No
Disclosure	+	o	+	o	–	+	No
Social	+	+	o	+	o	+	No
Solve	+	–	+	+	+	o	No
Ask	+	–	+	+	+	+	No
Handover	+	–	+	–	–	+	Yes
Persuasion	–	+	+	+	o	+	Yes
Assessment rating scale: + (suitable), o (adaptions necessary), – (not usable)							

Table 3: Assessment of Recovery Strategies

Handover: When analyzing literature involving breakdowns of CAs, we discovered that in the existing research, the handover from the CA to the human is extremely rarely discussed in the context of conversational breakdown recovery, although handover and characteristics are a current research interest (e.g., Poser et al. 2021 investigate CA-human handover and time delay during the handover process). This seems rather surprising since many real-world implementations of CAs frequently delegate the conversation to a human employee if the conversation breaks down and the CA cannot help. For example, many internet services that employ CAs for customer interaction to leverage the strengths of this technology (e.g., around-the-clock availability and low maintenance cost) have their CAs set up in the way that if the CA fails, the option to contact a real person is provided. In the context of our breakdown recovery strategies, this can be seen as a combination of existing recovery strategies extended by a potentially novel handover strategery. To hand over the conversation to a real human, the CA will first have to acknowledge the failure, thus using the confirmation strategy. Next, the CA will try to salvage the situation by suggesting a handover and hence using a solve strategy to find a solution for the interacting user. Therefore, our conceptual overview of six breakdown recovery strategies may be expanded by a seventh handover strategy. However, we chose not to do so since our focus was to derive strategies from literature, and we did not find any relevant literature on this subject. Nevertheless, we want to highlight this interaction as a potential new recovery strategy future research avenues may confirm. Although we believe that a CA should be designed to solve problems and user inquiries on its own, even if there are mis- or non-understandings, we believe it would be useful to investigate when exactly a handover to a human is required. For instance, as a final fallback strategy for when all other strategies fail. In this sense, critical situations could be identified in which recovery strategies are ineffective, necessitating the use of humans. It would be interesting to investigate whether this undermines trust in the CA and its competence in this regard. It would also be interesting to learn whether the user or customer is still satisfied with the overall experience of the interaction afterward or if this impacts their attitude toward CAs. To investigate this, we propose that it should be studied empirically in a real-world setting. An initial starting point could be the examination of influencing factors on the perceived usefulness. For instance, one could examine the time delay until the handover has been conducted and its impact on perceived usefulness.

User Group: We also discovered that the literature makes little distinction between external and internal users. This raises whether the breakdown strategies for internal users, such as employees who contact the chatbot, should be designed similarly to the breakdown strategies for external users. Moreover, many studies rely on students for experimentation which is an understandable choice from a practical perspective. However, the transferability to a real-world context and application may be questioned. Additionally, some studies use very small samples of about 10 test subjects for their empirical evaluation. With such a small sample, the empirical results may be questioned as well. While you must consider reputational damage or dissatisfied customers when dealing with external users, the costs associated with internal breakdowns are largely unknown. In this regard, it would be interesting to see what disadvantages conversational breakdowns currently cause for internal users and whether the associated costs, such as time lost due to CA not understanding you, can also be quantified. However, not only is the distinction between internal and external important but so is the distinction between user groups. The question here is whether different breakdown strategies for different user groups are required to avoid the "one strategy fits all" approach. There will almost certainly not be a single breakdown recovery strategy that all user groups fully grasp. Thereby, to meet the demand for a good experience, several breakdown recovery strategies would have to be developed and designed to include all possible cultures and backgrounds. The agents' ability to adapt to the user is therefore required in this case. With ever-improving technology, implementing this adaptability will only be a matter of time. In this sense, we can envision a series of studies on the personalization or individualization of CAs to determine which recovery strategy is best suited to what user groups.

Persuasive System Design: Another potential lead to extend the capabilities of breakdown recovery is the inclusion of persuasive design elements like clues, hints, or nudges in the CAs design; potentially even design elements that transform human-agent interaction into a gameful or playful experience. Such persuasive design elements are design modifications and mechanics that may encourage the user to behave in a specific way (Fogg 1998), although CAs on their own already can be seen as persuasive social actors (Fogg 2003). Nevertheless, designing CAs with persuasive elements is a current research interest and may prove a valuable addition to strategies for conversational breakdown recovery. For example, the CA may motivate the user to stay engaged in the interaction and provide additional information to help the CA recover from a potential breakdown and thus keep the conversation alive (Benner et al. 2021).

Overall, these three leads may prove as potentially valuable extensions to support theoretical contributions on conversational breakdown strategies. Additionally, while analyzing the relevant literature, we found that many studies only investigate whether a recovery mechanism is effective or not. However, the **degree of effectiveness** remains unclear, especially in comparison to other existing recovery strategies. An investigation of quantifiable effects of conversational breakdown strategies and combinations thereof, as we have observed strategies to be combined frequently, may prove another valuable lead for future research. We want to encourage fellow researchers and practitioners to pursue these leads and contribute to this important topic for human-agent interaction. To provide a starting point for future research avenues, we present the following research questions researchers may use:

Q1: How can handover be used for conversational breakdown recovery?
Q2: What is the distinction between internal and external users in human-agent interaction?
Q3: How can persuasive system design be used for conversational breakdown recovery?
Q4: What is the quantifiable effectiveness of each conversational breakdown recovery strategy?

Conclusion

Our goal was to address the state of art concerning conversational breakdown recovery strategies when interacting with CAs. Therefore, we conducted a systematic literature analysis across several databases and organized our findings in table 1. In total, we provide a two-fold contribution to research and practice with our research paper. Firstly, by conducting a systematic literature analysis, we describe the current status of the research on conversational breakdown and recovery strategies thereof. Simultaneously, our review covers the current state-of-the-art recovery strategies for conversational breakdowns. We describe what strategies are used. Additionally, we derive generalized categories for the recovery strategies that all have their unique characteristics. Secondly, we shed light on the issues that persist in the research area of conversational agents, conversational breakdowns, and recovery strategies. We observe that the community is still faced with a low level of applicability and transparency of conversational recovery strategies. This circumstance becomes very clear as some authors do not specify the employed recovery strategies in detail or what and how they address (e.g., Frampton and Lemon 2008; Komatani et al. 2009; Rieser and Lemon 2011). Consequently, we encourage future research to position and formulate recovery strategies. This bears the potential to foster transparency when interacting with CAs and minimizing distrust as well as dissatisfaction. We also anticipate doing so will provide more guidance for other CA designers and developers when instantiating respective recovery strategies into CAs.

Despite us following established guidelines and research methodology to rigorously analyze the identified literature on conversation breakdowns when interacting with CAs, our paper does not come without limitations. First, the scope of this SLA cannot claim to be fully exhaustive, as we did not cover all existing databases and potentially excluded some relevant literature by year, for example. Second, analogous to the selection of databases, our selection of keywords may not cover all possible areas of research that cover recovery strategies, although we addressed this limitation by rigorously conducting a forward and backward search process. However, we intended to reach a representative coverage of literature in the domain of CAs by applying a rigorous research method for searching and analyzing the papers. Third, the presented recovery strategies of CAs are rigorously derived from the prior analyzed research on conversation breakdowns. The categorization, however, may not be fully objective, as we had to make decisions on how to categorize our findings. Thus, a certain residual level of uncertainty and subjectivity remains as this process involves individual human judgment.

As we analyzed the literature and our findings, we noticed a pattern present in CAs and their breakdown recovery strategies. Many CAs use information or ask strategies for breakdown recovery (e.g., Gnewuch et al. 2017; Litman et al. 2006); some even combine both strategies (e.g., Bohus and Rudnicky 2005; Dzikovska et al. 2010; Kocaballi et al. 2020). These two strategies have a certain persuasive element in the way they function. For example, CAs can tell the users that providing more information can improve and prevent a conversational breakdown, thus nudging the user to comply with the CAs preventive mechanisms. As such, digital nudging (Thaler and Sunstein 2009) can be used to implement breakdown strategies with small design features and persuade the user to help the CA prevent a breakdown. This theme is currently being recognized by other authors as well, as they take up the beneficial properties of persuasiveness in CAs (Benner et al. 2021). In total, we provided an overview of existing research on recovery strategies for conversational breakdowns and possible future research directions on this topic with a hint towards persuasive features. We hope our contribution will help researchers and practitioners to design breakdown resilient CAs and continue research on this important topic.

Acknowledgments

We thank the European Regional Development Fund (grant No. 20007305) and the Swiss National Science Foundation (grant No. 100013_192718) for funding parts of this research. The fourth author acknowledges funding from the Basic Research Fund (GFF) of the University of St. Gallen.

References

Ashktorab, Z., Jain, M., Liao, Q. V., and Weisz, J. D. 2019. "Resilient Chatbots," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, S. Brewster, G. Fitzpatrick, A. Cox and V. Kostakos (eds.), Glasgow, Scotland Uk. 04.05.2019 - 09.05.2019, New York, New York, USA: ACM Press, pp. 1-12.

Benner, D., Schöbel, S., and Janson, A. 2021. "Exploring the State-of-the-Art of Persuasive Design for Smart Personal Assistants," *Wirtschaftsinformatik* (16).

Benner, D., Schöbel, S., and Janson, A. 2021. It Is Only for Your Own Good, or Is It? Ethical Considerations for Designing Ethically Conscious Persuasive Information Systems, presented at the Americas Conference on Information Systems (AMCIS), Association for Information Systems

Bickmore, T., Trinh, H., Asadi, R., and Olafsson, S. 2018. "Safety First: Conversational Agents for Health Care," in Studies in conversational UX design, R. J. Moore, M. H. Szymanski, R. Arar and G.-J. Ren (eds.), Cham: Springer, pp. 33-57.

Bitner, M. J., Brown, S. W., and Meuter, M. L. 2000. "Technology Infusion in Service Encounters," *Journal of the Academy of Marketing Science* (28:1), pp. 138-149.

Bittner, E., Oeste-Reiß, S., and Leimeister, J. M. 2019. "Where is the Bot in our Team? Toward a Taxonomy of Design Option Combinations for Conversational Agents in Collaborative Work," in

Proceedings of the 52nd Hawaii International Conference on System Sciences, T. Bui (ed.), Hawaii International Conference on System Sciences.
Black, H. 2018. *What is Conversation Design and How to Design Your Chatbot. Medium*. https://medium.com/swlh/what-is-conversation-design-and-how-to-design-your-chatbot-3754f04ab1e7. Accessed 26 February 2021.
Bohus, D., and Rudnicky, A. I. 2005. "Error handling in the RavenClaw dialog management framework," in *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing - HLT '05*, R. J. Mooney (ed.), Vancouver, British Columbia, Canada. 06.10.2005 - 08.10.2005, Morristown, NJ, USA: Association for Computational Linguistics, pp. 225-232.
Brandtzaeg, P. B., and Folstad, A. 2018. "Chatbots: changing user needs and motivations," *interactions* (25:5), pp. 38-43.
Brandtzaeg, P. B., and Følstad, A. 2017. "Why People Use Chatbots," in *Internet science*: *4th international conference, INSCI 2017, Thessaloniki, Greece, November 22-24, 2017 : proceedings*, Y. Kompatsiaris (ed.), Cham: Springer, pp. 377-392.
Bulyko, I., Kirchhoff, K., Ostendorf, M., and Goldberg, J. 2005. "Error-correction detection and response generation in a spoken dialogue system," *Speech Communication* (45), pp. 271-288.
Chakrabarti, C., and Luger, G. F. 2015. "Artificial conversations for customer service chatter bots: Architecture, algorithms, and evaluation metrics," *Expert Systems with Applications* (42:20), pp. 6878-6897.
Clark, L., Munteanu, C., Wade, V., Cowan, B. R., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., Spillane, B., Gilmartin, E., and Murad, C. 2019. "What Makes a Good Conversation?" in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, S. Brewster, G. Fitzpatrick, A. Cox and V. Kostakos (eds.), Glasgow, Scotland Uk. 04.05.2019 - 09.05.2019, New York, New York, USA: ACM Press, pp. 1-12.
Cooper, H. M. 1988. "Organizing knowledge syntheses: A taxonomy of literature reviews," *Knowledge in Society* (1:1), pp. 104-126.
Cooper, H. M. 1998. *Synthesizing research*: *A guide for literature reviews*, Thousand Oaks, Cal.: Sage Publ.
Dzikovska, M. O., Callaway, C. B., Farrow, E., Moore, J. D., Steinhauser, N., and Campbell, G. 2009. "Dealing with Interpretation Errors in Tutorial Dialogue,"
Dzikovska, M. O., Moore, J. D., Steinhauser, N., and Campbell, G. 2010. "The Impact of Interpretation Problems on Tutorial Dialogue,"
Engelhardt, S., Hansson, E., and Leite, I. 2017. "Better Faulty than Sorry: Investigating Social Recovery Strategies to Minimize the Impact of Failure in Human-Robot Interaction," *Workshop on Conversational Interruptions in Human-Agent Interactions at IVA 2017*.
Fast, E., Chen, B., Mendelsohn, J., Bassen, J., and Bernstein, M. S. 2018. "Iris: A Conversational Agent for Complex Tasks," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: Association for Computing Machinery, pp. 1-12.
Feine, J., Gnewuch, U., Morana, S., and Maedche, A. 2019. "A Taxonomy of Social Cues for Conversational Agents," *International Journal of Human-Computer Studies* (132), pp. 138-161.
Fettke, P. 2006. "State-of-the-Art des State-of-the-Art," *WIRTSCHAFTSINFORMATIK* (48:4), p. 431f.
Fogg, B. J. 1998. "Persuasive computers," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, M. E. Atwood (ed.), Los Angeles, California, United States. 4/18/1998 - 4/23/1998, New York, NY: ACM Press/Addison-Wesley Publishing Co, pp. 225-232.
Fogg, B. J. 2003. "Computers as persuasive social actors," in *Persuasive Technology*, Elsevier, pp. 89-120.
Folstad, A., and Brandtzaeg, P. B. 2017. "Chatbots and the new world of HCI," *interactions* (24:4), pp. 38-42.
Frampton, M., and Lemon, O. 2008. "Using Dialogue Acts to Learn Better Repair Strategies for Spoken Dialogue Systems," *IEEE International Conference on Acoustics, Speech and Signal Processing*.
Gnewuch, U., Morana, S., and Maedche, A. 2017. "Towards Designing Cooperative and Social Conversational Agents for Customer Service," *International Conference on Information Systems (ICIS)*.
Griol, D., Carbó, J., and Molina, J. M. 2013. "An Automatic Dialog Simulation Technique To Develop And Evaluate Interactive Conversational Agents," *Applied Artificial Intelligence* (27:9), pp. 759-780.
Gunkel, D. J. 2016. "Computational Interpersonal Communication: Communication Studies and Spoken Dialogue Systems," *communication+1* (5:1).

Holz, T., Dragone, M., and O'Hare, G. M. P. 2009. "Where Robots and Virtual Agents Meet," *International Journal of Social Robotics* (1:1), pp. 83-93.

Janssen, A., Passlick, J., Rodríguez Cardona, D., and Breitner, M. H. 2020. "Virtual Assistance in Any Context," *Business & Information Systems Engineering* (62:3), pp. 211-225.

Keyser, A. de, Köcher, S., Alkire, L., Verbeeck, C., and Kandampully, J. 2019. "Frontline Service Technology infusion: conceptual archetypes and future research directions," *Journal of Service Management* (30:1), pp. 156-183f.

Kocaballi, A. B., Coiera, E., and Berkovsky, S. 2020. "Revisiting Habitability in Conversational Systems," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, R. Bernhaupt (ed.), Honolulu HI USA. 25 04 2020 30 04 2020, New York,NY,United States: Association for Computing Machinery, pp. 1-8.

Komatani, K., Ikeda, S., Fukubayashi, Y., Ogata, T., and Okuno, H. G. 2009. "Ranking help message candidates based on robust grammar verification results and utterance history in spoken dialogue systems," in *Proceedings of the SIGDIAL 2009 Conference on The 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue - SIGDIAL '09*, P. G. T. Healey, R. Pieraccini, D. Byron, S. Young and M. Purver (eds.), London, United Kingdom. 11.09.2009 - 12.09.2009, Morristown, NJ, USA: Association for Computational Linguistics, pp. 314-321.

Komatani, K., and Okuno, H. G. 2010. "Online Error Detection of Barge-In Utterances by Using Individual Users' Utterance Histories in Spoken Dialogue System,"

Komatani, K., and Rudnicky, A. I. 2009. "Predicting barge-in utterance errors by using implicitly supervised ASR accuracy and barge-in rate per user," in *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers on - ACL-IJCNLP '09*, Unknown (ed.), Suntec, Singapore. 04.08.2009 - 04.08.2009, Morristown, NJ, USA: Association for Computational Linguistics, p. 89.

Komatsu, S., and Sasayama, M. 2019. "Speech Error Detection depending on Linguistic Units," in *Proceedings of the 2019 3rd International Conference on Natural Language Processing and Information Retrieval,* Tokushima Japan, New York, NY, USA: ACM, pp. 75-79.

Kontogiorgos, D., Pereira, A., Sahindal, B., van Waveren, S., and Gustafson, J. 2020a. "Behavioural Responses to Robot Conversational Failures," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, T. Belpaeme (ed.), Cambridge United Kingdom. 23 03 2020 26 03 2020, New York,NY,United States: Association for Computing Machinery, pp. 53-62.

Kontogiorgos, D., van Waveren, S., Wallberg, O., Pereira, A., Leite, I., and Gustafson, J. 2020b. "Embodiment Effects in Interactions with Failing Robots," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, R. Bernhaupt, F. '. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguey, P. Bjørn, S. Zhao, B. P. Samson and R. Kocielnik (eds.), Honolulu HI USA. 25 04 2020 30 04 2020, [S.l.]: Association for Computing Machinery, pp. 1-14.

Laumer, S., Maier, C., and Gubler, F. 2019. "CHATBOT ACCEPTANCE IN HEALTHCARE: EXPLAINING USER ADOPTION OF CONVERSATIONAL AGENTS FOR DISEASE DIAGNOSIS," *Research Papers*.

Lee, C., Jung, S., Lee, D., and Lee, G. G. 2007. "Example-Based Error Recovery Strategy for Spoken Dialog System," *2007 IEEE Workshop on Automatic Speech Recognition and Understanding*.

Lee, K., Ryu, S., Seo, P. H., Kim, S., and Lee, G. G. 2014. "Grammatical Error Corrections Based on Learner Comprehension Model in Oral Conversation," *2014 IEEE Spoken Language Technology Workshop (SLT 2014)*.

Lee, S., Lee, D., Hooshyar, D., Jo, J., and Lim, H. 2020. "Integrating breakdown detection into dialogue systems to improve knowledge management: encoding temporal utterances with memory attention," *Information Technology and Management* (21:1), pp. 51-59.

Lee, Y., Park, S., and Suk, H.-J. 2018. "Applying Social Strategies for Breakdown Situations of Conversational Agents: A Case Study using Forewarning and Apology," *Korean Society for Emotion and Sensibility* (21:1), pp. 59-70.

Li, T. J.-J., Chen, J., Xia, H., Mitchell, T. M., and Myers, B. A. 2020. "Multi-Modal Repairs of Conversational Breakdowns in Task-Oriented Dialogs," in *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, S. Iqbal (ed.), Virtual Event USA. 20 10 2020 23 10 2020, New York,NY,United States: Association for Computing Machinery, pp. 1094-1107.

Li, T. J.-J., Radensky, M., Jia, J., Singarajah, K., Mitchell, T. M., and Myers, B. A. 2019. "PUMICE: A Multi-Modal Agent that Learns Concepts and Conditionals from Natural Language and Demonstrations," in *Proceedings of the 32nd Annual ACM Symposium on User Interface Software*

and Technology, F. Guimbretière, M. Bernstein and K. Reinecke (eds.), New Orleans LA USA, New York, NY, USA: ACM, pp. 577-589.
Lieberman, H. 1997. "Autonomous interface agents," in *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, S. Pemberton (ed.), Atlanta Georgia USA. 22 03 1997 27 03 1997, New York, NY: ACM, pp. 67-74.
Litman, D., Hirschberg, J., and Swerts, M. 2006. "Characterizing and Predicting Corrections in Spoken Dialogue Systems," *Computational Linguistics* (32:3), pp. 417-438.
Luger, E., and Sellen, A. 2016. ""Like Having a Really Bad PA"," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, J. Kaye, A. Druin, C. Lampe, D. Morris and J. P. Hourcade (eds.), Santa Clara, California, USA. 07.05.2016 - 12.05.2016, New York, New York, USA: ACM Press, pp. 5286-5297.
Marge, M., and Rudnicky, A. I. 2019. "Miscommunication Detection and Recovery in Situated Human–Robot Dialogue," *ACM Transactions on Interactive Intelligent Systems* (9:1), pp. 1-40.
McTear, M. F. 2017. "The Rise of the Conversational Interface: A New Kid on the Block?" in *Future and Emerging Trends in Language Technology. Machine Learning and Big Data*: *Second International Workshop, FETLT 2016, Seville, Spain, November 30 -December 2, 2016, Revised Selected Papers*, J. F. Quesada, F.-J. Martín Mateos and T. López Soto (eds.), Cham: Springer International Publishing, pp. 38-49.
Meredith, J. 2017. "Analysing technological affordances of online interactions using conversation analysis," *Journal of Pragmatics* (115), pp. 42-55f.
Moore, R., and Arar, R. 2019. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*.
Mozafari, N., Weiger, W., and Hammerschmidt, M. 2020. "Resolving the Chatbot Disclosure Dilemma: Leveraging Selective Self-Presentation to Mitigate the Negative Effect of Chatbot Disclosure," *Proceedings of the Hawaii International Conference on System Sciences*.
Myers, C., Furqan, A., Nebolsky, J., Caro, K., and Zhu, J. 2018. "Patterns for How Users Overcome Obstacles in Voice User Interfaces," in *Engage with CHI*: *CHI 2018 : proceedings of the 2018 CHI Conference on Human Factors in Computing Systems : April 21 -26, 2018, Montréal, QC, Canada*, R. Mandryk and M. Hancock (eds.), Montreal QC, Canada. 4/21/2018 - 4/26/2018, New York, New York: The Association for Computing Machinery, pp. 1-7.
Nordheim, C. B., Følstad, A., and Bjørkli, C. A. 2019. "An Initial Model of Trust in Chatbots for Customer Service—Findings from a Questionnaire Study," *Interacting with Computers* (31:3), pp. 317-335.
Porcheron, M., Fischer, J. E., Reeves, S., and Sharples, S. 2018. "Voice Interfaces in Everyday Life," in *Engage with CHI*: *CHI 2018 : proceedings of the 2018 CHI Conference on Human Factors in Computing Systems : April 21 -26, 2018, Montréal, QC, Canada*, R. Mandryk and M. Hancock (eds.), Montreal QC, Canada. 4/21/2018 - 4/26/2018, New York, New York: The Association for Computing Machinery, pp. 1-12.
Poser, M., Singh, S., and Bittner, E. 2021. "Hybrid Service Recovery: Design for Seamless Inquiry Handovers between Conversational Agents and Human Service Agents," *Hawaii International Conference on System Sciences 2021*.
Qiu, L., and Benbasat, I. 2009. "Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems," *Journal of Management Information Systems* (25:4), pp. 145-182.
Radziwill, N. M., and Benton, M. C. 2017. "Evaluating Quality of Chatbots and Intelligent Conversational Agents," *Software Quality Professional*.
Rieser, V., and Lemon, O. 2011. "Learning and Evaluation of Dialogue Strategies for New Applications: Empirical Methods for Optimization from Small Data Sets," *Computational Linguistics* (37:1), pp. 153-196.
Sagawa, H., Mitamura, T., and Nyberg, E. 2004. "Correction grammars for error handling in a speech dialog system," in *Proceedings of HLT-NAACL 2004: Short Papers on XX - HLT-NAACL '04*, J. Hirschberg, S. Dumais, D. Marcu and S. Roukos (eds.), Boston, Massachusetts. 02.05.2004 - 07.05.2004, Morristown, NJ, USA: Association for Computational Linguistics, pp. 61-64.
Schegloff, E. A., Jefferson, G., and Sacks, H. 1977. "The preference for self-correction in the organization of repair in conversation," *Language* (53:2), pp. 361-382.
Schuetzler, R. M., Grimes, G. M., and Scott Giboney, J. 2020. "The impact of chatbot conversational skill on engagement and perceived humanness," *Journal of Management Information Systems* (37:3), pp. 875-900.

Shankar, V., Smith, A. K., and Rangaswamy, A. 2003. "Customer satisfaction and loyalty in online and offline environments," *International Journal of Research in Marketing* (20:2), pp. 153-175.

Skarbez, A. Kotranza, F. P. Brooks, B. Lok, and M. C. Whitton 2011. "An initial exploration of conversational errors as a novel method for evaluating virtual human experiences," in *2011 IEEE Virtual Reality Conference*, pp. 243-244

Stoyanchev, S., and Johnston, M. 2015. "Localized Error Detection for Targeted Clarification in a Virtual Assistant," *2015 IEEE International Conference on Acoustics, Speech and Signal Processing*.

Stoyanchev, S., Salletmayr, P., Yang, J., and Hirschberg, J. 2012. "Localized Detection of Speech Recognition Errors," *IEEE Spoken Language Technology Workshop (SLT 2012)*, pp. 25-30.

Stoyanchev, S., and Stent, A. 2009. "Predicting concept types in user corrections in dialog," in *Proceedings of the 2nd Workshop on Semantic Representation of Spoken Language - SRSL '09*, M. Alcantara-Pla and T. Declerck (eds.), Athens, Greece. 30.03.2009 - 30.03.2009, Morristown, NJ, USA: Association for Computational Linguistics, pp. 42-49.

Sutcliffe, A. 2009. "Designing for User Engagement: Aesthetic and Attractive User Interfaces," *Synthesis Lectures on Human-Centered Informatics* (2:1), pp. 1-55.

Takayama, J., Nomoto, E., and Arase, Y. 2019. “Dialogue breakdown detection robust to variations in annotators and dialogue systems,” *Computer Speech & Language* (54), pp. 31-43

Thaler, R. H., and Sunstein, C. R. 2009. *Nudge*: *Improving decisions about health, wealth, and happiness*, New York, NY: Penguin.

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., and Plattfaut, R. 2015. "Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research," *Communications of the Association for Information Systems* (37).

Watson, R. T., and Webster, J. 2020. "Analysing the past to prepare for the future: Writing a literature review a roadmap for release 2.0," *Journal of Decision Systems* (29:3), pp. 129-147.

Winkler, R., Hobert, S., Salovaara, A., Söllner, M., and Leimeister, J. M. 2020. "Sara, the Lecturer: Improving Learning in Online Education with a Scaffolding-Based Conversational Agent," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, R. Bernhaupt, F. '. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguey, P. Bjørn, S. Zhao, B. P. Samson and R. Kocielnik (eds.), Honolulu HI USA. 25 04 2020 30 04 2020, [S.l.]: Association for Computing Machinery, pp. 1-14.

Wintersberger, P., Klotz, T., and Riener, A. 2020. "Tell Me More: Transparency and Time-Fillers to Optimize Chatbots' Waiting Time Experience," in *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, D. Lamas, H. Sarapuu, I. Šmorgun and G. Berget (eds.), Tallinn Estonia. 25 10 2020 29 10 2020, New York,NY,United States: Association for Computing Machinery, pp. 1-6.

Woodward, J., McFadden, Z., Shiver, N., Ben-hayon, A., Yip, J. C., and Anthony, L. 2018. "Using Co-Design to Examine How Children Conceptualize Intelligent Interfaces," in *Engage with CHI*: *CHI 2018 : proceedings of the 2018 CHI Conference on Human Factors in Computing Systems : April 21 - 26, 2018, Montréal, QC, Canada*, R. Mandryk and M. Hancock (eds.), Montreal QC, Canada. 4/21/2018 - 4/26/2018, New York, New York: The Association for Computing Machinery, pp. 1-14.

Xiao, Z., Zhou, M. X., and Fu, W.-T. 2019. "Who Should Be My Teammates: Using a Conversational Agent to Understand Individuals and Help Teaming," in *Proceedings of the 24th International Conference on Intelligent User Interfaces*, New York, NY, USA: Association for Computing Machinery, pp. 437-447.

Zierau, N., Elshan, E., Visini, C., and Janson, A. 2020. A Review of the Empirical Literature on Conversational Agents and Future Research Directions. Proceedings of the 41st ICIS.