“Please Connect Me to a Specialist”: Scrutinising ‘Recipient Design’ in Interaction with an Artificial Conversational Agent

Iuliia Avgustis[1*[0000-0003-1026-7852]], Aleksandr Shirokov[2[0000-0002-5374-5466]],and Netta Iivari[1[0000-0002-7420-2890]]

[1] University of Oulu, Oulu, Finland
{iuliia.avgustis, netta.iivari}@oulu.fi
[2] Rutgers University, New Brunswick, USA
aleksandr.shirokov@rutgers.edu

Abstract. This paper explores how callers formulate information enquiries for an artificial conversational agent in a call centre and compares it with the way enquiries are addressed to human operators of the same call centre. It includes 60 call recordings with human operators and 103 call recordings with the artificial conversational agent, transcribed and analysed using the method of Conversation Analysis. We show that people formulate and reformulate their enquiries differently to an artificial agent, even though the goal in both cases is to get an answer to the same enquiry. When talking to the artificial conversational agent, callers produce short enquiries, similar to web searches. When connected to human operators, callers formulate longer enquiries which include many details. By analysing the differences in the way callers formulate their enquiries to robots and human operators, we show what callers expect artificial conversational agents to process. These expectations affect the way the enquiry is formulated and, as a result, operators and artificial agents encounter different types of problems they have to repair to understand the question correctly and find an answer to it. Our findings have interesting implications for Human Computer Interaction both in terms of “robot-recipient design” and “user-recipient design”.

Keywords: Conversational agent, speech interface, call centre, conversation analysis, recipient design

1 Introduction

Currently, there is a growing use and continuous development of conversational agents, voice chatbots, voice user interfaces (VUI) and personal assistant devices. They have attracted wide scholarly attention in the field of human-computer interaction (HCI). Recent studies in this field have focused particularly on interaction with advanced conversational agents (Siri, Amazon Echo, Google Home, etc.) in everyday settings, mostly by families in the context of their homes. The studies have emphasized the intricacies involved in embedding the interaction with the agent into the everyday activities and

conversational settings of the participants (see e.g. [10, 36, 37, 48]). The studies reveal interaction with the agent is being associated with entertainment in many cases, people laughing and being amused [16, 29, 36, 37, 48]. Our study contributes to the stream of research on conversational agents, but directs attention to a different kind of conversational setting: it explores interaction with a conversational agent in a task-oriented institutional setting, in which interaction with the agents differs dramatically from interaction with more advanced conversational agents; the agents follow a very limited, strict set of rules and scenarios of interaction, the task-oriented human participants expecting effectiveness and efficiency above all from the encounters. This kind of less advanced conversational agents are widely used in different kinds of domains such as in survey interviews or service encounters, while they receive clearly less attention in research on interaction with conversational interfaces.

Our study aims to fill in this knowledge gap and investigates features of interaction with less advanced but still human-like conversational agents in task-oriented institutional interaction. The research question of this study is "How formulations and reformulations of enquiries addressed to artificial agents differ from those addressed to human operators?" The element of comparison is important for understanding the specifics of conversations with artificial agents. Therefore, the first section of the analysis is devoted to the comparison of initial formulations, and in the second section we will discuss the difference between the way callers reformulate inquiries to human operators and artificial agents. We consider important to investigate this question, as artificial agents in institutional settings are usually designed based on similar types of interaction with human operators who also follow certain rules. However, people talk differently to non-human agents, they design their enquiries in a different way based on their expectations, and as a result, the input for the agent is different. If we understand how callers talk to artificial agents, we can design artificial agents in institutional setting that will better understand the enquiries. Specifically, we are interested in a caller's initial formulation of a question or an enquiry and its reformulation if the first attempt does not provide the expected answer from an agent. Some differences between the ways people communicate with humans and artificial agents have already been mentioned in the setting of everyday online conversations with chatbots [23]. However, there are still relatively few studies comparing differences between the way people communicate with human operators and voice bots.

This study provides an important opportunity to advance the understanding of the organized structure of interaction with artificial agents (i.e. telephone robots) in telephone enquiry services. It also offers some insights into callers' expectations, reasoning process and challenges that arise during this type of institutional interaction. Telephone robots are now widely used in call centres, as they dramatically reduce human operators' workload. However, callers still often prefer talking to a human operator, as robots cannot always comprehend the enquiry. Understanding features of enquiries formulated for artificial agents can therefore clarify what robots should be able to process. We conclude this study with ideas on how the quality of interaction with artificial agents in institutional calls can be improved.

2 Related research

2.1 Interaction with conversational agents in different settings

Interaction with human-like dialogue systems (systems that aim to imitate human conversation [15]) has become a part of our everyday social life. One can find plenty of HCI research addressing artificial conversational agents and VUIs over the past decades, considering solutions such as search engines, help systems, intelligent personal assistants, or robots. Currently, one can find an increasing HCI interest on conversational agents [3, 4, 5, 10, 13, 14, 16, 18, 29, 30, 36, 37, 48], the studies being inspired by the emergence of advanced conversational agents (Siri, Amazon Echo, Google Home, etc.) in everyday settings. These studies examine how families embed artificial conversational agents into the everyday activities of the home, how they are used as resources in interaction, and how the agents are fitted into the sequential organization of the talk, identifying how the agent is addressed and how its responses are handled [16, 29, 36, 37]. Besides home and family use, the studies also explore conversational agents in other entertainment-oriented contexts, such as in museums [4].

We wish to direct attention back to the task-oriented institutional contexts with high requirements for effectiveness and efficiency. Recently, this type of context has received less attention in HCI. However, we maintain research interest is warranted in this type of context, as conversational interfaces have become widely spread and used within. Studies also show interaction problems still prevail in interaction with conversational agents in task-oriented contexts (e.g. [10, 30, 49]). This study focuses on human interaction with an artificial conversational agent in a task-oriented intuitional setting of a call centre. Technology oriented studies addressing call centre or telephone service VUIs or can be found in the computing field [1, 9, 26, 34] as well as studies addressing the user experience and usability of such services and VUIs [28, 43, 46]; however, these studies do not focus in detail on human interaction with a conversational agent or address the influence of the institutional setting.

Calls to telephone helplines and various other telephone services are a type of institutional talk. To understand such talk, we turn to Conversation Analysis, within which it has been shown that the telephone enquiry service, as an institutional setting, imposes certain rules and restrictions on social interaction [21]. This type of interaction has features that distinguish it from ordinary conversation: (a) participants are oriented to specific goals (e.g. obtaining information); (b) the interaction has constraints on what is treated as appropriate contributions to achieving this goal; (c) it is associated with frameworks and procedures that are specific to the institutional context (e.g., digital or paper documentation, etc.) [22]. Human operators in call centres are required to follow written instructions when they introduce themselves, ask clarifying questions and give answers to callers. The rules, however, do not fully determine the practices. They can also be ignored or made explicit and discussed by participants to the communication. Interaction with an artificial agent in a telephone service has a similar structure and it is organized mainly through question-answer adjacent pairs. However, the range of possible response utterances in interaction with an artificial agent is much more limited, and if a caller does not follow the rules, s/he cannot get an answer to her/his enquiry.

Human operator can depart from the instructions if necessary, i.e. when communication trouble arises. One could assume it is relatively straightforward to design a well-functional conversational agent in this type of setting, while in this study we will show a lot of trouble may emerge in interaction, showing problems in recipient design and repair. Next, we will elaborate on these concepts.

2.2 Conversation with a conversational agent

The wide use of various voice technologies makes relevant the debate of the 1990s about the possibility of human-machine conversation [32]. One of the key figures of this debate is Graham Button who stated that only a “simulacrum of conversation” [8], i.e. an imitation as opposed to a real conversation between humans, is possible in interaction with machines. A similar a priori distinction between “human-human” and “human-machine” interaction is found in new research on HCI [38], in which conversation between people is contrasted with intercourse between humans and conversational agents. However, some researchers argue that interaction between a human and a machine can be considered a conversation and analysed as such as long as there are sense-making procedures done by the human [50].

Studies show that conversation with artificial agents differs from conversations with humans. For example, the language in conversations with voice interfaces is found to be command-like and abrupt [19]. Speakers adapt to the system’s abilities when they talk to an artificial agent. The concept of recipient design refers to this phenomenon: as a concept used in CA studies, it refers to the tuning of the utterances for different addressees and is performed by people in their everyday interactions [39]. In this paper, we will address its specific category - robot-recipient deign, where the recipient is an artificial agent. Interaction with a non-human recipient raises additional issues for the interlocutor, and we will analyse them in the context of formulations and reformulations of the enquiry in institutional calls. Various aspects of "robot-recipient design" will be identified and discussed throughout the analysis (how callers adapt their initial enquiries when connected to a robot, and how they treat troubles in these conversations). The notion of "robot-recipient design" and its design implications will be more thoroughly addressed in the discussion section.

Prior HCI studies have examined recipient design also from another angle: recipient design has been introduced as a VUI principle in IBM Conversational UX guidelines [16, 33], in which recipient design refers to the adaptation of the artificial agent to the actions of user and her/his knowledge. There is already discussion on request and response design in HCI research along these lines. Studies empirically examine how users formulate requests to an agent and react to its responses [16, 36, 48], the studies also highlighting that designers of artificial agents need to consider what they intend users to say to the agent and how to design resourceful responses for users [10, 16].

Repair has also received attention in HCI research on conversational agents. Different kind of repair strategies have been discussed [3, 37, 48], both from the viewpoint of the agent and the user. The topic of conversation repair has originally been introduced by Schegloff, Jefferson, and Sacks [41], defined as practices that address troubles in speaking, hearing, and understanding. The goal of repair is to maintain or restore

intersubjectivity and to move interaction forward [40]. Initially, this phenomenon was described using English-language data, but since then repair has been analysed in various languages [42], including Russian [7]. Repair can be initiated and carried out by one actor, but it can also be initiated by one actor and later be carried out by another actor [27]. Conversation analysts found a structural "preference" for self-repair [41] and a tendency that most often the repair is initiated by the speaker of the "trouble-source" [27]. This highlights one specific problem of interaction with a "simple" call centre agent: even though they can initiate the repair themselves (and thus signal the existence of a problem), they are usually not capable of performing a repair (and thus actually solving the problem). Thus, in such interaction, a significant part of the work of carrying out repair remains with the user.

In this study we address a task-oriented institutional context of a call centre and particularly focus on recipient design and repair in human-agent interaction within: we examine how humans adapt to the agents' assumed abilities when talking with the agent and how the agent has been designed to adapt to users' abilities (or not). We discuss the repair strategies employed in interaction, showing pertinent problems within. It has already been found that repairs are more frequent in task-oriented conversations than in "ordinary" dialogues [12]. A novel angle in this study is comparison of human-agent interaction with human-human interaction in the same setting. This reveals that people talk differently to artificial agents in institutional settings from the beginning of interaction, and therefore, artificial agents should not be designed based on interaction with human operators.

We join the HCI discussion on naturalness of conversation with artificial agents in institutional settings: we show users engaging in 'robotic' interaction with the agent, but also the agent trying to engage to 'human-like' interaction with users. Especially in institutional, task-oriented settings we should be wary of the illusion of natural conversation: the existing literature has already pointed out ([48], see also [5]) that the better the agent works from the users' perspective, the more naturally also the users tend to speak, which increases the risk of trouble in interaction due to giving users an illusion of natural conversation, users remaining or becoming unaware of the agent's limited capabilities. We highlight these concerns as especially acute in task oriented institutional settings.

3 Research design

3.1 Data

The data for this study were collected at one major Russian city's call centre that answers citizens' questions considering different administrative issues such as official documents and working hours of public institutions. The artificial conversational agent of this centre was created by this organizations' employees to reduce the workload. It has been programmed to answer every fourth call during a given test period. In calls with the artificial agent, callers are informed that they will talk to a robot (this is the term that we will also use in the paper). The data for this research were provided by the

call centre representatives for the analysis and publication purposes. All ethical guidelines were followed.

The collected data include 60 call recordings with humans and 103 call recordings with the artificial agent. Collected audio recordings of naturally occurring conversations with human and robot operators were transcribed and analysed using the method of Conversation Analysis (CA). In this paper, we will present examples of the most frequent types of formulations and reformulations in interaction with artificial agents and human operators, which will allow us to reveal the main differences. In calls with the artificial agent, callers are first informed that they will speak to a robot and then instructed to state the question and speak after the tone. All callers are also informed about the audio recording of the call in this welcome message. After this message the recording, as well as most of the transcripts presented in this paper, start and the caller hears the tone (line 7 in the transcripts presented further).

Human operators in this call centre follow strict instructions and use specified wordings when answering a question. As the artificial operator is representing this institution, and as its answer is considered to be an official response from the governmental structure, it has to follow certain rules of interaction as well. For this reason, its developers have made the decision to create a simple rule-based conversational agent instead of a generative model conversational agent. They gave the robot a set of predefined responses and wrote strict scripts that this robot is supposed to follow. The robot chooses what to say next based on the keywords and connections between these words that it finds in the recognized speech of the caller. For example, if the robot recognizes "readiness" and "social card" in the enquiry of the caller, it asks to indicate the social card application number. The robot's voice is a female voice, and its quality is rather good; however, pauses between some words and phrases are sometimes either too short or too long, and intonations are often inappropriate, so while the voice sounds natural, the speech does not. Despite that, our collection also includes several cases where people confuse the robot with an actual human operator and address it as such. These cases are not presented in this paper due to their infrequency.

Conversations with the robot are ended either by providing information or by transferring the call to a human operator if the robot does not identify relevant keywords or does not have a predefined answer to a particular question. We will provide examples for each of these cases. Both human and robot operators are in the same organization, the same telephone enquiry service, and they both talk to the same type of callers. Therefore, the collected data provides a good opportunity for making a comparison between interaction with an artificial agent and interaction with a human operator.

3.2 Method

To analyse the data, we use the method of Conversation Analysis (CA), which was originally established as a method for studying everyday conversation and interaction but was later used to study other forms of interaction, including interaction between humans and machines [20, 44, 45]. CA has already proved to be useful in the field of human-computer interaction and computer-supported cooperative work [2, 11]. It is also often used to study interaction with artificial agents [6, 16, 24, 31, 32, 35, 36].

CA is a qualitative data driven method that draws the researcher's attention to details of naturally occurring interaction. This focused attention allows the researcher to identify methods that members use to accomplish their activities and make these activities accountable to others (e.g. systematic patterns for asking a question or for formulating an enquiry). To analyse the methods of performing actions, a researcher needs to analyse the sequence of actions in one specific conversation. Conversation Analysis does not require analysis of a large number of conversations, but it makes necessary a close attention to the details of interaction in specific situations. The main question therefore is "how" participants carry out a specific social action - instead of "how often" or "why". However, usually conversation analysts investigate more than one case, as the analysis and comparison of different cases may reveal how a specific method works in different situations and what variations of this method exist.

To capture as many details of the interaction as possible, Gail Jefferson's system of transcription [25] is used by conversation analysts. Creating a Jeffersonian transcript is a necessary research step that allows the researcher to see how the conversation actually unfolds. Pauses, changes in speed of talk, intonations and other features of the talk are transcribed to ensure a better understanding of actions accomplished in and by the talk. The following abbreviations are used in the following transcripts: C – caller, R – robot operator, H – human operator. All the personal data (names, application numbers etc.) as well as the name of the service have been changed to preserve anonymity.

We will first focus on methods of formulating an enquiry, and then analyse how initial formulations are repaired by the participants of interaction in various situations.

4 Findings

4.1 Formulating an enquiry

If the robot answers the call, the caller hears the welcome message and receives the instruction in one sentence: "please clearly formulate your question and speak after the tone". In interactions with human operators, people are able to produce or formulate a social action and expect that this action will be correctly recognized by the interactional partner. In the interaction with the telephone robot, the caller is faced with the question of what to say and how to say it so that the robot would recognize what has been said [38], i.e. how to design an enquiry for a robot. This is partially due to the fact that the caller has to formulate an enquiry before he or she actually hears the robot and estimates what it is capable of. However, it is also a characteristic of the subsequent interaction, as such interactional agents are not capable of producing actions that simultaneously show what these actions are doing [38].

The instruction – "please clearly formulate your question" does not actually help the caller to understand how to interact with the robot. What does "clearly" mean? Does this imply "in detail" or "briefly" or with respect to articulation? Clearly for whom? The problem of this uncertainty for a caller is indicated by the fact that after the welcome message and the first beep a caller often stays quiet instead of asking a question, as seen in the line 8 in the following example. As we mentioned earlier, all transcripts

of interaction with the robot operator start from the line 7, as the caller first hears the introduction message, and only then the audio recording begins, so even though it is the beginning of the interaction for the robot, the caller has previously listened to a short pre-recorded instruction by a human on stating the question after the tone.

Example 1

```
07  R  ((tone))
08     (2.5)
09  C  .hh allo  [mne °by°   ]
       .hh hello [I'd like to]
10  R             [pozhalu-   ]
                  [plea-      ]
11     (2.1)
12  C  .h soedinite menja pozhalujsta so ↑spetsialistom.
       .h please connect me to a ↑specialist.
```

The gap (line 8) is typical only for interaction with a robot operator, and there is no such gap in the interactions with a human operator. After the greeting from a human operator and the question “how can I help you?”, a caller immediately begins to speak. We can assume that the initial formulation would be different if the robot would also start with "how can I help you" instead of "please clearly formulate your question and speak after the tone", but this requires a different set of data. Long periods of silence instead of a response are known to be one of the indicators of interactional trouble and they are usually minimized in interaction between humans [40]. In conversation with a robot operator, the average length of a gap is 2.4 seconds, and it is longer than 1 second in 78% of the cases we have analysed. The robot is programmed to say the phrase “please speak up” (line 10 – here the cut-off) when there is a long gap. Similar straightforward methods are rarely found in communication between humans, even in an institutional context. As can be seen from this transcript, the method itself can lead to an overlap of utterances (lines 9-10), although the robot has been programmed to stop speaking when there is an overlap. This interaction is not successful so far and it is continued by the caller’s request to talk to a human operator, who is considered to be a specialist as opposed to the robot (line 12). It is also noteworthy that the caller uses the word of politeness ("please") in this case. These words are rare in this type of interaction, and these enquiries often lead to a transfer to a human operator.

The instruction “clearly formulate your question” can be interpreted in different ways by callers. For example, there are several cases when the caller begins to talk in detail about her/his problem and talks not only about what s/he needs to find out, but also why s/he needs to do so. Often after these requests, the robot uses such replies as “maybe I misunderstood you” or “I'm sorry, I cannot understand” and asks again to “clearly formulate the question”. In most of the cases, however, callers make their enquiries short and similar to a web-based search request. In addition, callers often make short pauses after each word and stretch the pronunciation of words (apparently trying to make their speech more understandable for the robot). This "robot-like" speech is one of the aspects of robot-recipient design. Formulating an enquiry as a command

similar to a search request leads to successful interaction with the robot, and people quite often make an enquiry this way (Example 2). In the data we have analyzed, 75% of enquiries addressed to a robot operator were formulated this way.

Example 2

```
07  R  ((tone))
08     (3.2)
09  C  uznat' o gotovnosti patenta.
       to get information on readiness of a patent.
```

In the second example we also see a long gap, after which an enquiry without any explanations or additional details is provided (line 9 in Example 2). It is further successfully read by the robot, so the caller receives the requested information. This way of formulating a request is specific for interacting with the robot. In cases when a caller is redirected from a robot operator to a human, the same request is formulated quite differently, as seen in the Example 3. The caller first speaks to a robot (lines 8-9 and omitted lines 10-37) and is later connected to the human operator, as interaction with the robot did not lead to the expected by the caller answer.

Example 3

```
07  R  ((tone))
08     (2.1)
09  C  xochu uznat' (0.4) zadolzhennost' po kvartplate.
       I'd like to know (0.4) rent arrears.
[...]
38  O  vy pozvonili v tsentr predostavlenija gosuslug Kontakt,
       you called to the service center Kontakt,
39     spetsialist Veselova Ol'ga dobryj den'.
       specialist Veselova Olga good afternoon.
40     (0.7)
41  C  zdravstvujte .hh >podskazhite pozhalujsta<
       hello .hh >please tell me<
42     a ja mogu vot po telefonu uzna:t' summu zadolzhennosti.=
       can I fi:nd out on the phone the amount of debt.=
43     =↑ja tak primerno znaju,
       =↑I know more or less,
44     no vot (1.2) xotela (1.7) ne znaju tam
       but well (1.2) I wanted (1.7) I don't know
45     .hh prosto s sentjabrja my ne platim i .hh a::: °e°,
       .hh it's just that we don't pay since September .hh u::hm °e°,
46     schitat' summu kotoraja mne prixodit na: >nu vot<,
       do I count the sum that comes to: >well<,
47     v pochtovyj ↑jaschik,
       to the ↑mailbox,
48     ili tam kakie-to eschë protsenty byvajut.
```

```
        or are there some other penalties.
49      >°ja prosto ne znaju°<
        >°I just don't know°<
```

In this case, even though the robot understands the request stated by the caller, it is unable to provide the caller the necessary information because of the lack of knowledge on the question, so the caller is redirected to a human operator. As seen from this transcript, the same caller formulates the request in a different way depending on the operator: a very short enquiry formulated to a robot (line 9) and a long, detailed enquiry formulated to a human operator (lines 41-49). Note also that the gap after operator's answer (line 40) is much shorter than the gap in interaction with the robot (line 8).

In the second formulation, in addition to the same enquiry, there is a greeting and words of politeness (line 41). Apparently, successful interaction with this artificial agent rarely involves greetings, introductions or other polite forms of conversation. Moreover, in the interaction with a human operator, the caller often gives the circumstances of the call, the reasons for the question to arise or details of the question (lines 45). The question can also be repeated twice during one turn to make it more clear for the human operator (lines 42 & 46-48). All of these elements are rarely present in the case of interaction with the robot operator, and all of them might cause additional problems for the robot. The formulation of a request as a command is a way to successfully interact with the robot and it is the exact way that most people formulate their enquiries.

We can see that the callers orient to the properties of the recipient in their turns-at-talk, and the "recipient design" of initial enquiries is different in the talk with the human and the robot, as the robot is simply introduced as such, and nothing is said about its capabilities. However, although enquiries are different, both the robot and the human operators can rarely answer the question after the first turn of the caller is over. Usually some additional work is needed: either by specifying the request or by clarifying certain details of it. In the next section we will concentrate on repair sequences that people perform to correct their initial enquiry and formulate a question in an adequate form.

4.2 Repair sequences

Repair situations are a useful focus for analysing what we call robot-recipient design. The type of repair solution shows a recipient (and an analyst) how the actor treats the trouble source (as a problem of speaking, hearing, or understanding) and can indicate the actor's assumptions about the recipient. We found a specific pattern of what repair solutions callers tend to propose and accordingly how they tend to interpret trouble sources in interaction with the artificial conversational agent. One of the situations in which a repair sequence typically occurs is when the robot (for an unknown reason) does not recognize speech fully or correctly, as in the following example.

Example 4

```
07  R  ((tone))
08     (1.5)
09  C  gotova ↑li sotsial'naja karta?
```

```
        ↑is the social card ready?
10      (4.1)
11  R   utochnite,
        specify,
12      gotovnost' (.) kakogo dokumenta vas interesuet
        readiness (.) of which document you are interested in
13      (0.7)
14  R   ((tone))=
15  C   =.h SOTSIAL'NA- SOTSIAL'NAJA KARTA
        =.h SOCIA- SOCIAL CARD
```

In line 9, the caller produces an enquiry as a simple question. Even though the enquiry is short and precise, the robot doesn't recognize the name of the document. In lines 11-12, the robot produces a repair initiation, which specifies the document's name/type as the trouble source. Next, there is a 0.7 gap (line 13) and a tone (line 14). The caller doesn't answer immediately but waits for the tone. This shows her orientation to the way the robot works. Right after the tone, the caller produces a repair solution (line 15). She uses the same name for defining the document but pronounces it louder. After this, the robot succeeds in correctly understanding the enquiry and continues to follow its script of actions.

In this case, the caller treats the trouble source as a problem of hearing (or speech recognition) since she doesn't change the document's name. However, the robot didn't specify the type of problem. This could be both a hearing problem and a formulation problem (the document's incorrect name). In such ambiguous situations, callers tend to interpret trouble sources as a problem of hearing. In their repair solution, they don't change the "content" but pronounce it differently, usually louder and with longer pauses between words. This is what we call a robot-like voice. Let us consider one more case in which a robot initiates a repair sequence.

Example 5

```
11  R   nazovite (.) pozhalujsta
        indicate (.) please
12      nomer zajavlenija na sotsial'nuju kartu
        the social card application number
13      (.)
14  R   govorite posle signala.
        speak after the tone.
15      (0.4)
16  R   ((tone))
17      (1.3)
18  C   znachit (0.2) PJAT' (0.2) DVA (0.2) ↑SHEST' (0.8) ODIN
        so (0.2) FIVE (0.2) TWO (0.2) ↑SIX (0.8) ONE
19      (0.3) ↓PJAT' (0.8) TRI (0.4) CHETYRE (0.3) DVA.
        (0.3) ↓FIVE (0.8) THREE (0.4) FOUR (0.3) TWO.
20      (4.1)
```

```
21  R  vozmozhno ja vas nepravil'no ponjala
       maybe I misunderstood you
22     (.)
23  R  nazovite nomer esche raz po odnoj tsifre
       indicate the number one digit after another
24     (.)
25  R  pozhalujsta govorite posle signala
       please speak after the tone
26     (0.4)
27  R  ((tone))
28     (0.8)
29  C  PJA:T' (0.8) DVA (0.4) SHEST' (0.8) ODIN
       FI:VE (0.8) TWO (0.4) SIX (0.8) ONE (0.6)
30     (0.6) PJAT' (0.5) TRI (0.8) CHETYRE (0.4) DVA.
       (0.6) FIVE (0.5) THREE (0.8) FOUR (0.4) TWO.
```

In lines 11-14, the robot asks for the application number. The caller provides it (lines 18-19), but the robot, for some reason, doesn't accept it. In lines 21-25, the robot initiates a repair by explicating the misunderstanding and giving more detailed instructions. In response, the caller pronounces the same number louder and makes longer pauses between words (lines 29-30). So again, the robot doesn't specify the type of problem, and the caller treats it as a hearing problem.

In the above-mentioned cases, the robot's demand for clarification ignores the fact that the caller has already mentioned all the necessary information. This type of questioning is seen quite often in talks with the robot, but it is rarely seen in interactions with human operators. The latter usually initiate the repair in another way. They propose a repair solution and request confirmation or rejection of the solution. Human operators have instructions they have to follow if the reason for the call is not established after the caller's first turn. To achieve clear understanding of the question, operators are supposed to use following phrases: "You mean that …", "That is, you want …", "So you need …", "You need to clarify …" etc. In most of the cases they also make a guess about the question and ask the caller if they understood them correctly. Here is an example:

Example 6

```
07  C  >dobryj vecher Marina, menja zovut Anna<,
       >good evening Marina, my name is Anna<,
08     ja xotela uznat',
       I want to know,
09     my:: s papoj privatiziruem kvarti:ru,
       me:: and father are privatizing the apa:rtment,
10     mne by interesno bylo uznat' ↑sro:ki
       I am interested in knowing the ↑te:rms
11     .hh vot s momenta kak my ↑podaem
       .hh from the moment we ↑apply
```

```
12      vse vot eti ↑dokumenty v ↓portal,
        all these ↑documents to the ↓portal,
13      cherez ↑kakoe vremja ona budet privatizirovana,
        after ↑what time it will be privatized,
        ((the operator starts to type and types until line 17))
14      my prosto ee ↑prodaem,
        we just ↑sell it,
15      poetomu mne by ukazat' etu ↑datu.
        so I would like to indicate the ↑date.
16      (2.8)
17  H   neposredstvenno ↑srok ispolnenija.
        (you mean) ↑document processing time.
18      ↓pravil'no          [°ja vas ponjala°?
        have I ↓correctly [°understood you°?
19  C                       [da >da da<
                            [yes >yes yes<
```

The initial enquiry is formulated through lines 8-15. The caller gives a reason for a call and describes the situation. The operator starts to look for the answer in the data base (line 13) before the caller finishes her turn, which might be the reason why she is able to quickly transform this enquiry to a simple question further. Even though the gap in line 16 is rather long for an interaction with a human operator, the caller as well as the analyst can hear the typing by the operator, which makes her actions accountable. This gap, therefore, is different from those we have seen in interactions with the robot. In lines 17-18, the operator initiates the repair and proposes the repair solution by reformulating the question. As the caller confirms that the enquiry formulated by the operator is correct, she will then get an answer from the operator. In most of the cases, as well as in the one presented here, human operators have to work with the details mentioned in the initial enquiry to help the caller to formulate a question. However, sometimes they also don't get enough details and have to lead the caller towards formulating an enquiry in multiple turns. This is done by phrases such as "yes", "I listen to you", "speak further" etc. In cases, where an operator is not sure about what was said, a guess is also usually made.

Robots are also sometimes faced with an enquiry which is either too detailed or not detailed enough. This could lead to the robot's incorrect (from a caller's point of view) answer as it understands the question by connecting keywords that are recognized during the initial enquiry. In the following example, words in the enquiry were successfully recognized, but not in the way the caller expected the enquiry to be recognized.

Example 7

```
07  R   ((tone))
08      (0.6)
09  C   .h >MFTS Danilovskogo rajona<
        .h >MFC of Danilovskogo district<
10      (4.2)
```

11	R	tsentr Gosuslug rajona Danilovskij
		the Gosuslug center of the Danilovskij district
12		naxoditsja po ↑adresu
		is located at the ↑address
13		(0.3)
14	C	.tch
15		(0.2)
16	R	ulitsa Xavskaja ↑dom dvadtsat' ↓shest'
		Xavskaja street ↑house twenty ↓six
[...]		
29	R	↑vam povtorit'?
		shall I repeat?
30		(0.6)
31	R	((tone))
32		(3.5)
33	C	>soedinite pozhalujsta< s MFTS Danilovskogo rajona.
		>please connect< to MFC of Danilovskogo district.

Although the formulation of an enquiry in the form of a command assumes a short statement, it must be accurate enough; otherwise, the robot can provide wrong type of information. In the above transcript, the caller learns to formulate the request while interacting with the robot. The first request (line 9) was read by a robot as a request for information about the Multifunctional Centre for Provision of State and Municipal Services (MFC) of a certain area – its location, working hours, etc. (lines 11-16 and omitted lines 17-28). However, as it becomes clear that the robot is going to give this information about the mentioned MFC, the caller produces ".tch" sound which can be interpreted as a sign of discontent. This is something a human operator could orient to and stop, but the robot continues. The caller does not attempt to interrupt the robot either and waits until the end of its turn to add new keywords to the initial enquiry (line 33).

The initial enquiry wouldn't be formulated to the operator in the same way, so it is not possible to make a comparison of a similar trouble source in a conversation with a human. Human operators usually get long and detailed enquiries that often include an account for the enquiry and description of the situation. Even though human operators also need to have an enquiry in a short and simple form to find an answer in their data base, callers tend to transfer this duty to the operator. In case there are not enough details, human operator indicates to the caller that there are not enough details by encouraging the caller to speak further. It is also possible that human operator can misunderstand the question and start to read the answer that is not relevant. In this case, unlike the robot, the human operator can be stopped by the caller, and the enquiry can be reformulated once again before the full answer is given by the operator.

In situations where the robot experiences speech recognition or keyword search failure that eventually leads to the repetition or reformulation of the initial enquiry by the caller, the robot does not use the same methods as human operators use when it needs to make its problem explicit. However, the callers manage to reformulate their enquiries based on their knowledge about how the robot works. In most of the cases, the callers

speak either slower or louder or make longer pauses. There are also different combinations of this "voice robotization". Differences in the users' understanding of the reason for interactional trouble can be seen in the way initial enquiries are reformulated. As the robot can only indicate that there was a trouble, the callers tend to treat it as a problem of hearing and repeat their enquiries louder and with longer pauses between words. When talking to human operators, callers rarely encounter troubles in hearing, so a comparison of similar cases is not possible. However, when other problems arise, operators provide a possible solution to the problem (see Example 6).

5 Discussion

5.1 Summary of the results

Although the number of artificial conversational agents in institutional settings is constantly growing, there are relatively few studies on how people actually interact with them. Investigating this aspect is important, as it reveals how callers design their utterances for an artificial agent to achieve the goal of institutional interaction. If we understand how people talk to an institutional agent, we can design an agent that will better deal with callers' enquiries, offer smoother and more efficient interaction and ultimately better customer and user experience. In institutional, task-oriented settings, pleasurable customer and user experiences are of utmost importance for continued service or product use and customer and user retention (e.g. [28, 43, 46]). In this paper, we focused on the first step in interaction with an artificial agent in a telephone enquiry service; the formulation of an information enquiry and its repair after the operator's reply.

Even though both human operators and robots follow instructions when talking to callers, they follow them differently, as the robot cannot deviate from the instructions. These differences can especially be seen in how repairs are initiated and carried out. However, as we have shown an initial enquiry is also formulated differently to a robot and a human, and accordingly different types of troubles in interaction arise. In the talk with the human operator, the request includes many details: reason for the call, description of the situation, etc. (see Example 3, 6). In general, requests to a human operator are longer. Therefore, to repair the caller's enquiry, the human operator shortens it and translates it into the form of a "clearly formulated question" to which the operator can find an answer in a knowledge base (see Example 6). A human operator, therefore, initiates and carries out the repair in most of the cases.

Despite the institutional character of the call which imposes certain rules on the interaction, people formulate and reformulate their enquiries differently to a robot, even though the goal in both cases is to get an answer to the same enquiry (see Example 3). Callers design their utterances based on their knowledge of the recipient, so when talking to a robot they orient to its (in)capabilities. They adapt their responses and perform "robot-recipient design" (as a form of a recipient design). When talking to the robot, people tend to produce simple and unambiguous enquiries (see Examples 2, 3, 4, 7) before giving the details in subsequent answers to robot's questions. In the case of the "robot-recipient design", enquiry is formulated as a command or a search request; hence, the potential difficulties – an incorrectly composed command and the need to

reformulate it as a way of carrying out the repair, or a lack of necessary information and the need to add it in the next turn. Repair is usually initiated by the robot in a case of technical or interactional trouble, but the caller is the one who is expected to provide the solution (see Examples 4, 5, 7).

5.2 Research and design implications

Findings of this study lead to several implications for HCI researchers, designers of conversational user experience, and developers of artificial conversational agents for institutional settings. Our findings have interesting implications both in terms of "robot-recipient design" and "user-recipient design".

This study introduces the notion of robot-recipient design as a novel form of recipient design for HCI research on artificial conversational agents. It highlights that information enquiries are produced differently for human operators and for an artificial agent. Types of repairs are also different in interaction with a human and a robot operator. Overall, this study maintains that humans orient to their assumption about an artificial agent's conversational competence and (in)capabilities, and they adjust their utterances to this knowledge. Our results offer a description of these adjustments.

Robot-recipient design has been acknowledged by HCI research to an extent. The use of a less rich language in interaction with robots has already been noted in previous research [23]. Short pauses after each word and stretch of the pronunciation of words have been found in prior studies on interaction with conversational interfaces [36]. However, this study explicates and makes this form of recipient design more visible, indicating also a new research area for the future: on users' (often implicit) assumptions about the robot-recipients' capabilities – there is a need to study their formation (how they form, when and where, based on what kind of experiences, expectations, bias, preknowledge, trajectories) as well as their evolution along people's encounters with different kinds of ever evolving intelligent systems and agents. Such research should determine how the assumptions are enacted in the actual conversation, i.e. in this case their initial enquiries and repairs, in different settings and with different systems. Similarly to the older HCI research stream of mental models, we should start examining what different kinds of people know, expect and assume in relation to artificial conversational agents' capabilities, acknowledging that very likely their understanding is not factual but may well be superstitious, outdated and biased. This, however, should be done based on the analysis of real conversations. We should also follow up the developments in technology and see how people's understanding is evolving and adapting along their encounters with ever more intelligent systems and agents.

We wish to direct HCI attention also to another form of recipient design: "user-recipient design", performed by designers and developers of artificial agents (see also [10, 16, 33]). Our study indicates that making the robot's capabilities more explicit to the caller might improve and speed up the interaction. This is especially relevant for artificial agents in institutional settings, as their main purpose is getting a task done fast and easy, e.g. through getting enough information to find an answer to an enquiry. In more entertainment-oriented settings people may tolerate or even enjoy more complicated interactions with a lot of reformulation and repair. Our data showed callers tended

to formulate their enquiries to human operators in a way that required additional work on the interpretation and reformulation. Both initial requests for an enquiry and requests for its repair should be formulated differently by an artificial agent to avoid the necessity of this additional work. Making explicit what a robot can understand (e.g. by giving callers an input example) will lead to more appropriately formulated information enquires. This will eventually improve the process of parsing, transcribing and processing of an enquiry by the robot. Robot's voice and way of speaking (pauses, intonations, etc.) should correspond to its capabilities, so that callers get a better idea about how they should formulate their enquiries. As we have shown, callers make their speech more robotic (and easier to parse) after they hear a reply from the robot. If the robot introduces itself instead of being introduced in a message pre-recorded by a human, callers can better evaluate robot's capabilities before formulating an initial enquiry. The robot's introduction should be carefully designed to include relevant information on its capabilities. We also suggest that the less natural robot's speech and voice sounds, the less expectations people have, which results in easier formulations for the robots (see also [5, 48]). This hypothesis, however, has to be investigated with a different set of data. Future research could also focus on the comparison of enquiries formulated for the robot by people who interact with it for the first time and by those who have had previous experience of talking to this agent. This could reveal the process of learning to talk to an artificial agent in institutional settings.

Overall, our study leads us to questions about the future of robots in institutional settings as well as in our everyday life: should robots learn to talk like humans (e.g. Google Duplex), or should humans learn to talk like robots; should robots in institutional setting be more natural or should they be more effective; should conversations with robots be merged with our everyday life or should they be recognized as related to technology use? Should our children be taught to human-human as well as human-robot interaction? How should robots introduce themselves to us to make us aware of their nature and capabilities? The HCI research community should be more engaged in reflection on these very broad issues, as human-robot interactions are becoming a prominent part of our everyday life. It is important to remember that not only developers perform "user-recipient design", i.e. design robot's utterances, but users also design their responses, i.e. perform "robot-recipient design". A question for a future study could be on the dynamics between robot-recipient and user-recipient design: "How does robot-recipient design change when the design of the robot changes?" One interesting implication of our study is also that to understand how a human would interact with a robot, it is not sufficient to analyse human interactions of the same type, as it has been done in previous studies [47]. We also wish to point out the significance of emotions and affective computing in relation to human interaction with robots. As we have an interaction-based approach of CA, our focus has been on interactional features and we have not tried to find expressions of emotional stances. However, emotions are an interesting topic for future studies in this context.

This study demonstrates the usefulness of CA in studies on interaction with artificial agents. It has already been known that humans talk to robots differently [17], but CA allows us to see what exactly those differences are [35]. The method focuses on micro level analysis of empirical data and is not strong in generating design implications.

Nevertheless, we show CA results can be used as a basis for design implications regarding specific studied phenomena. We acknowledge the idea of robot-recipient design should be developed further through empirical investigations; we have taken only an initial step towards developing this approach. As an initial set of design guidelines, we propose the following: 1) Always consider the target user group(s) and their background knowledge and experiences with conversational agents (of different kind): those will inform their robot-recipient design and should guide the user-recipient design; 2) design the agent so that it makes its (in)capabilities visible for the user as early as possible (consider carefully how natural the interaction should seem for users); 3) design the agent so that it guides the user towards appropriate robot-recipient design (offering help in the formulation of initial enquiries as well as in repair). Naturally, these guidelines need to be developed towards a more practical design approach. Important is also to acknowledge that there are different types of robot-recipient design (e.g. for virtual assistant-recipient design, telephone robot-recipient design, etc.); we have empirically explored only one particular case; future studies are needed in this respect.

6 Conclusion

To conclude, in the talk with a robot, callers produce "robot-robot" conversation as they start to produce "robot-like" responses themselves, while in the talk with human operators, callers tend to shift the task of formulating the "clearly formulated question" to the operator. In situations when a repair is necessary, operators also have a bigger role: the robot indicates only the trouble source, while operators also provide a possible solution to the trouble. Interaction with the robot has a particular conversational character, as people constantly orient to their knowledge on robot's (in)capabilities. This raises a question regarding whether robots should be designed based on interaction with humans. There are always discussions in the field of development of conversational agents on how to make a machine speak like a person, but if a robot needs to get an enquiry in a particular form that is easy for him to read and find an answer to it is actually the question of how to make robot more "robot-like" that becomes relevant. It is especially relevant in the design of artificial agents for institutional settings, as the main goal of these agents is to understand an enquiry and provide a relevant response effectively and efficiently. Having a more "robot-like" agent will decrease callers' expectations regarding robots' capabilities and lead to an easier to parse enquiry, which will be correctly interpreted by the robot. This will eventually lead to the relevant response and satisfactory experience for people who use the service.

Acknowledgements. Iuliia Avgustis received financial support from the research project "Smart Communication" (2018-2022, Eudaimonia Institute, University of Oulu). We would like to thank members of the EMCA_Ru Research Group (A. Korbut, A. Maximova, K. Popova, A. Reiniuk, N. Belov) in collaboration with whom data for this research were collected and transcribed and who provided valuable insights during data sessions. Special thanks should be given to Florence Oloff for constructive comments and useful critiques, as well as to Gary David for his insightful feedback.

References

1. Acomb, K., Bloom, J., Dayanidhi, K., Hunter, P., Krogh, P., Levin, E., Pieraccini, R.: Technical support dialog systems: issues, problems, and solutions. In: Proceedings of the Workshop on Bridging the Gap: Academic and Industrial Research in Dialog Technologies. pp. 25–31. Association for Computational Linguistics, USA (2007).
2. Albert, S., Housley, W., Stokoe, E.: In case of emergency, order pizza: an urgent case of action formation and recognition. Presented at the Proceedings of the 1st International Conference on Conversational User Interfaces August 22 (2019). https://doi.org/10.1145/3342775.3342800.
3. Ashktorab, Z., Jain, M., Liao, Q.V., Weisz, J.D.: Resilient Chatbots: Repair Strategy Preferences for Conversational Breakdowns. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–12. Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3290605.3300484.
4. Barth, F., Candello, H., Cavalin, P., Pinhanez, C.: Intentions, Meanings, and Whys: Designing Content for Voice-based Conversational Museum Guides. In: Proceedings of the 2nd Conference on Conversational User Interfaces. pp. 1–8. Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3405755.3406128.
5. Barzilai, G., Rampino, L.: Just a Natural Talk? The Rise of Intelligent Personal Assistants and the (Hidden) Legacy of Ubiquitous Computing. In: Design, User Experience, and Usability. Design for Contemporary Interactive Environments. pp. 18–39. Springer International Publishing (2020).
6. Bennett, G.A.: Conversational Style: Beyond the Nuts and Bolts of Conversation. In: Moore, R.J., Szymanski, M.H., Arar, R., and Ren, G.-J. (eds.) Studies in Conversational UX Design. pp. 161–180. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-319-95579-7_8.
7. Bolden, G.B.: Negotiating Understanding in "Intercultural Moments" in Immigrant Family Interactions. Communication Monographs. 81, 208–238 (2014). https://doi.org/10.1080/03637751.2014.902983.
8. Button, G.: Going Up a Blind Alley: Conflating Conversation Analysis and Computational Modelling. In: Computers and conversation. pp. 67–90. Academic Press (1990).
9. Cena, F., Torre, I.: Increasing performances and personalization in the interaction with a call center system. In: Proceedings of the 9th international conference on Intelligent user interfaces. pp. 226–228. Association for Computing Machinery, New York, NY, USA (2004). https://doi.org/10.1145/964442.964487.
10. Clark, L., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., Spillane, B., Gilmartin, E., Murad, C., Munteanu, C., Wade, V., Cowan, B.R.: What Makes a Good Conversation? Challenges in Designing Truly Conversational Agents. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–12. Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3290605.3300705.
11. Clinkenbeard, M.: Multimodal Conversation Analysis and Usability Studies: Exploring Human-technology Interactions in Multiparty Contexts. Communication Design Quarterly. 6, 103–113 (2018). https://doi.org/10.1145/3282665.3282675.
12. Colman, M., Healey, P.: The Distribution of Repair in Dialogue. In: Proceedings of the Annual Meeting of the Cognitive Science Society. 33, (2011).
13. Dev, J., Camp, L.J.: User Engagement with Chatbots: A Discursive Psychology Approach. In: Proceedings of the 2nd Conference on Conversational User Interfaces. pp. 1–4. Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3405755.3406165.

14. Dubiel, M., Cervone, A., Riccardi, G.: Inquisitive mind: a conversational news companion. Presented at the Proceedings of the 1st International Conference on Conversational User Interfaces August 22 (2019). https://doi.org/10.1145/3342775.3342802.
15. Edlund, J., Gustafson, J., Heldner, M., Hjalmarsson, A.: Towards human-like spoken dialogue systems. Speech Communication. 50, 630–645 (2008). https://doi.org/10.1016/j.specom.2008.04.002.
16. Fischer, J.E., Reeves, S., Porcheron, M., Sikveland, R.O.: Progressivity for voice interface design. Presented at the Proceedings of the 1st International Conference on Conversational User Interfaces August 22 (2019). https://doi.org/10.1145/3342775.3342788.
17. Fischer, K.: How People Talk to Computers, Robots, and Other Artificial Communication Partners. Proceedings of the Workshop Hansewissenschaftskolleg, Delmenhorst April 21-23, (2006).
18. Følstad, A., Skjuve, M.: Chatbots for customer service: user experience and motivation. Presented at the Proceedings of the 1st International Conference on Conversational User Interfaces August 22 (2019). https://doi.org/10.1145/3342775.3342784.
19. Harris, R.A.: Voice Interaction Design: Crafting the New Conversational Speech Systems. Elsevier (2005).
20. Heath, C., Luff, P.: Technology in Action. Cambridge University Press (2000).
21. Heritage, J.: Conversation Analysis and Institutional Talk. In: Handbook of Language and Social Interaction. Lawrence Erlbaum Associates, Inc. (2005).
22. Heritage, J., Clayman, S.: Talk in Action: Interactions, Identities, and Institutions. Wiley-Blackwell (2010).
23. Hill, J., Ford, W.R., Farreras, I.G.: Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. Computers in Human Behavior. 49, 245–250 (2015).
24. Hutchby, I.: Conversation and Technology: From the Telephone to Internet. Polity (2001).
25. Jefferson, G.: Glossary of transcript symbols with an introduction. In: Lerner, G.H. (ed.) Conversation Analysis: Studies from the First Generation. pp. 13–34. John Benjamins Publishing Company (2004).
26. Jung, H., Kim, H., Ha, J.-W.: Understanding Differences between Heavy Users and Light Users in Difficulties with Voice User Interfaces. In: Proceedings of the 2nd Conference on Conversational User Interfaces. pp. 1–4. Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3405755.3406170.
27. Kitzinger, C.: Repair. In: The Handbook of Conversation Analysis. Wiley-Blackwell (2013).
28. Koca, A., Brombacher, A.C.: Extracting "broken expectations" from call center records: *why* and *how*. In: CHI '08 Extended Abstracts on Human Factors in Computing Systems. pp. 2985–2990. Association for Computing Machinery, New York, NY, USA (2008). https://doi.org/10.1145/1358628.1358795.
29. Lahoual, D., Frejus, M.: When Users Assist the Voice Assistants: From Supervision to Failure Resolution. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–8. Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3290607.3299053.
30. Li, T.J.-J., Chen, J., Xia, H., Mitchell, T.M., Myers, B.A.: Multi-Modal Repairs of Conversational Breakdowns in Task-Oriented Dialogs. In: Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology. pp. 1094–1107. Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3379337.3415820.
31. Li, T.J.-J., Labutov, I., Myers, B.A., Azaria, A., Rudnicky, A.I., Mitchell, T.M.: Teaching Agents When They Fail: End User Development in Goal-Oriented Conversational Agents.

In: Moore, R.J., Szymanski, M.H., Arar, R., and Ren, G.-J. (eds.) Studies in Conversational UX Design. pp. 119–137. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-319-95579-7_6.
32. Luff, P., Gilbert, N., Frohlich, D. eds: Computers and conversation. Academic Press, Bodmin, Cornwall (1990).
33. Moore, B., Arar, R.: Talk meets technology: Conversation design guidelines, https://conversational-ux.mybluemix.net/design/conversational-ux/.
34. Patel, N., Chittamuru, D., Jain, A., Dave, P., Parikh, T.S.: Avaaj Otalo: a field study of an interactive voice forum for small farmers in rural India. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 733–742. Association for Computing Machinery, New York, NY, USA (2010).
35. Pelikan, H.R.M., Broth, M.: Why That Nao?: How Humans Adapt to a Conventional Humanoid Robot in Taking Turns-at-Talk. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. pp. 4921–4932. ACM, New York, NY, USA (2016). https://doi.org/10.1145/2858036.2858478.
36. Porcheron, M., Fischer, J.E., Reeves, S., Sharples, S.: Voice Interfaces in Everyday Life. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. p. 640:1-640:12. ACM, New York, NY, USA (2018). https://doi.org/10.1145/3173574.3174214.
37. Porcheron, M., Fischer, J.E., Sharples, S.: "Do Animals Have Accents?": Talking with Agents in Multi-Party Conversation. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. pp. 207–219. Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/2998181.2998298.
38. Reeves, S.: Some conversational challenges of talking with machines. Presented at the Talking with Conversational Agents in Collaborative Action, Workshop at the 20th ACM conference on Computer-Supported Cooperative Work and Social Computing (CSCW '17), Portland, Oregon, USA February 25 (2017).
39. Sacks, H., Schegloff, E.A., Jefferson, G.: A Simplest Systematics for the Organization of Turn-Taking for Conversation. Language. 50, 696–735 (1974). https://doi.org/10.2307/412243.
40. Schegloff, E.A.: Schegloff, Emanuel A. Sequence organization in interaction: A primer in conversation analysis. Cambridge university press, New York (2007).
41. Schegloff, E.A., Jefferson, G., Sacks, H.: The Preference for Self-Correction in the Organization of Repair in Conversation. Language. 53, 361–382 (1977). https://doi.org/10.2307/413107.
42. Sidnell, J.: Comparative Studies in Conversation Analysis. Annual Review of Anthropology. 36, 229–244 (2007). https://doi.org/10.1146/annurev.anthro.36.081406.094313.
43. Sporka, A., Franc, J., Riccardi, G.: Can machines call people?: user experience while answering telephone calls initiated by machine. In: CHI Extended Abstracts (2009). https://doi.org/10.1145/1520340.1520545.
44. Suchman, L.A.: Plans and Situated Actions: The Problem of Human-machine Communication. Cambridge University Press, New York, NY, USA (1987).
45. Suchman, L.A.: Human-Machine Reconfigurations: Plans and Situated Actions. Cambridge University Press (2007).
46. Suhm, B., Peterson, P.: Call browser: a system to improve the caller experience by analyzing live calls end-to-end. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1313–1322. Association for Computing Machinery, New York, NY, USA (2009). https://doi.org/10.1145/1518701.1518899.

47. Szymanski, M.H., Moore, R.J.: Adapting to Customer Initiative: Insights from Human Service Encounters. In: Moore, R.J., Szymanski, M.H., Arar, R., and Ren, G.-J. (eds.) Studies in Conversational UX Design. pp. 19–32. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-319-95579-7_2.
48. Velkovska, J., Zouinar, M.: The illusion of natural conversation: interacting with smart assistants in home settings. Presented at the Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI'18), Montreal, Canada (2018).
49. Vtyurina, A., Savenkov, D., Agichtein, E., Clarke, C.L.A.: Exploring Conversational Search With Humans, Assistants, and Wizards. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. pp. 2187–2193. Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3027063.3053175.
50. Wooffitt, R., Fraser, N.M., Gilbert, N., McGlashan, S.: Humans, Computers and Wizards: Analysing Human (Simulated) Computer Interaction. Routledge (2013).