What Makes a Good Conversation? Challenges in Designing Truly Conversational Agents

Leigh Clark
leigh.clark@ucd.ie
University College Dublin, Ireland

Nadia Pantidi
konstantia.pantidi@ucc.ie
University College Cork, Ireland

Orla Cooney
orla.cooney@ucdconnect.ie
University College Dublin, Ireland

Philip Doyle
pdoyle@voysis.com
Voysis Ltd., Ireland

Diego Garaialde
Justin Edwards
diego.garaialde@ucdconnect.ie
justin.edwards@ucdconnect.ie
University College Dublin, Ireland

Brendan Spillane
Emer Gilmartin
brendan.spillane@adaptcentre.ie
gilmare@tcd.ie
Trinity College Dublin, Ireland

Christine Murad
cmurad@taglab.ca
University of Toronto, Canada

Cosmin Munteanu
cosmin@taglab.ca
University of Toronto Mississauga,
Canada

Vincent Wade
vincent.wade@adaptcentre.ie
Trinity College Dublin, Ireland

Benjamin R. Cowan
benjamin.cowan@ucd.ie
University College Dublin, Ireland

Abstract

Conversational agents promise conversational interaction but fail to deliver. Efforts often emulate functional rules from human speech, without considering key characteristics that conversation must encapsulate. Given its potential in supporting long-term human-agent relationships, it is paramount that HCI focuses efforts on delivering this promise. We aim to understand what people value in conversation and how this should manifest in agents. Findings from a series of semi-structured interviews show people make a clear dichotomy between social and functional roles of conversation, emphasising the long-term dynamics of bond and trust along with the importance of context and relationship stage in the types of conversations they have. People fundamentally questioned the need for bond and common ground in agent communication, shifting to more utilitarian definitions of conversational qualities. Drawing on these findings

we discuss key challenges for conversational agent design, most notably the need to redefine the design parameters for conversational agent interaction.

CCS Concepts

• **Human-centered computing** → **User studies**; **Natural language interfaces**; *HCI theory, concepts and models.*

Keywords

Conversational Agents, Speech HCI, Spoken Dialogue Systems, Voice User Interface Design, Interviews

1 Introduction

The proliferation of Intelligent Personal Assistants (IPAs) (e.g. Siri and Amazon Alexa) and the importance of speech in embodied agent technologies (e.g. robotics and embodied conversational agents (ECAs)) make human-agent conversation a critical topic of study. The IPA market alone is predicted to reach $9 billion by the early 2020s [14]. Speaking to agents currently falls short of conversational dialogue [37, 39, 42],

instead offering interactions that are constrained and task oriented. Although not necessarily desirable for all interactions, more human-like conversational ability could be important in supporting long-term human-agent interaction. This is especially true in contexts where social interaction and bond are important for delivering particular services (e.g. health, social care or education).

To address the current deficit in conversational ability, developers have largely focused on imbuing agents with more human-like conversational qualities such as social talk [4, 52] and humour [21]. The general aim is to emulate elements, and thus the qualities, of human conversation [27]. This approach, while important for generating conversation-like behaviour, risks losing focus on the subjective qualities that make conversation what it is. It also does not take into account the qualities users feel should manifest in conversational agent interactions, and how they should be implemented. Our work contributes to filling this gap by identifying: 1) what characteristics people see as important in conversation, and 2) how these vary when applied to conversations with artificial agents.

Through semi-structured interviews, we found that people clearly identified social and transactional roles of conversation, with almost universal focus on transactional purposes when discussing agent conversations. Participants emphasized similar characteristics as important in both human and agent-based conversation (e.g. mutual understanding, trustworthiness, active listenership and humour). Yet they were operationalized very differently, being discussed in almost purely functional terms for agent conversations. Participants described how conversational form and purpose vary with friends, strangers and acquaintances, emphasising the importance of conversation in long term relationship development. In reflecting on conversations with agents, participants' descriptions echoed human conversations with strangers or casual acquaintances. They fundamentally questioned the need to develop relationships with agents, seeing the agent as a tool rather than a dynamic social entity. Our findings are novel in that they contribute important user led insight to the growing debate around developing conversational ability in spoken agents. Based on these findings we emphasise the need to redefine conversational agent design parameters, based on the distinct differences between people's experiences of human-human and human-agent conversations.

2 Related Work

What is Conversation?

Spoken conversation is defined as: "any interactive spoken exchange between two or more people" [13, p.11] Human spoken conversation serves many purposes. These are broadly classified as transactional (task-based) or social (interactional) [11, 25, 51]. Transactional conversation pursues a practical goal, often fulfilled during the course of one interaction. In these types of exchanges, both interlocutors know what the goal of the dialog is. They have different clearly-defined roles, and success is measured by the achievement of the transaction's purpose. The aim of more social conversation is not to complete a task as such, but to build, maintain and strengthen positive relations with one or more interlocutors [23, 36]. Social conversation ranges from small talk and social greetings to longer interactions. Examples include talk between friends, office chat, or brief discussions between strangers. This type of social conversation can help develop common ground [16], trust and rapport between interlocutors [15]. Although transactional and social talk serve different purposes, they often overlap in natural conversation [15].

Critical to a conversation is the opening of a channel by an interlocutor, with a commitment from the other to engage, each then using the dialogue to co-construct meaning and converge on agreement [22]. Particularly in task oriented dialogue, this may lead to the proposition of an action to be completed or a transaction to take place.

Conversational Agents

The recent popularity of IPA devices like Amazon Echo and Google Home has generated considerable interest in the HCI community [40, 42]. Recent emphasis in HCI has been on understanding user experiences of interactions with IPAs. This work shows that people tend to engage in limited and clearly delineated task-based conversations with IPAs. These include checking the weather, setting reminders, calling and messaging, playing music, launching other applications, information seeking, and interacting with Internet of Things (IoT) devices [2, 20, 35, 38, 54]. Although these devices promise much through their implied humanness [20], they fall short of the reflexive and adaptive interactivity that occurs in most human-human conversation [42]. Instead interactions revolve around brief question/answer or user-instructions/system-confirmation dialogues [42]. Even where elements of social talk are implemented this seldom transcends the constraint of a single question and response. This is a far cry from the elaborative, contextual social talk seen in human conversation [28]. The highly functional and utilitarian nature of IPA interaction has led to suggestions that 'conversation' is a poor description of the current interaction experience [37, 42]. This deficit has motivated a number of efforts to improve the conversational capabilities in such agents [26, 41]. Recent work has looked to imbue systems with the ability to engage in social talk [52] and humour [21].

Although the current focus in HCI has been on commercial IPA experience, research on ECAs and social robotics have explored the user experience of more conversational interactions with agents. A number of these agents have been developed for contexts including health [6, 7], elderly care [5, 55], education [50], customer service [3] and workplace [29] contexts. ECA research [3] showed that users trusted an agent more when small talk was used and found the interaction more human like, although this varied by the user's personality type and level of ECA embodiment [4]. Similarly, in social robotics, conversational interaction is described as important to long term rapport building and use [53]. The accommodation of social conversations is proposed as a critical factor in developing trust in a social agent [34]. Work by Sabelli et al. [47] found that people were open to speaking about personal matters with a conversational robot, using it as an emotional outlet that reduced feelings of loneliness.

Research Aims

Social aspects of conversation may not be useful in all agent interactions. However, they may be beneficial in situations where the agent needs to build trust or rapport or needs to engage in frequent long-term user interaction. Emulating the structures and rules that govern human conversation [27] help generate conversation-like abilities, but they do not guarantee that the important subjective characteristics of conversation are preserved. Our work aims to identify: 1) what characteristics people see as important in conversation, and 2) how these vary when applied to conversations with artificial agents. Our goal is not to explore conversational properties, which are well known, but to emphasise the innate differences between the two types of conversations. We achieve this by presenting the results of semi-structured interviews on people's understandings and expectations of conversations with people and agents, and what contexts they expect speech-capable agents to be most suited to in the future.

3   Method

Participants

Seventeen participants (M=9, F=8) were recruited from a university community via internal email. In line with best practice, participants were recruited until saturation had occurred. Demographic data collected as part of the interview showed participants had a mean age of 31.1 years (SD=8.58) and consisted of students and staff across a wide range of schools. The majority of participants were native English speakers (64.7%), with the remaining participants speaking English to a near-native level (35.3%). Most participants rated their expertise with technology as advanced (41.2%) or intermediate (35.3%), with fewer describing themselves as experts

(23.5%). The majority of participants indicated they had previously used voice-based assistants (76.5%), although of this majority (58.8%) said they used them very infrequently. Participants reported interacting with Siri more than other assistants (56.3%), followed by Alexa (31.3%) and Google Assistant (31.3%). Participants were provided with a €10 honorarium in exchange for taking part.

Procedure and Data Analysis

After receiving university ethics clearance, interviews were conducted over a period of three weeks, lasting an average of 40 minutes, and audio-recorded with the participants' consent. The interviews were semi-structured, providing us with the flexibility to adjust questioning based on participant responses. Each interview covered four central topics: 1) Important characteristics, purposes and experiences of conversations with different types of people e.g. friends, acquaintances, strangers; 2) General attitudes towards conversations with agents; 3) Reflection on the important characteristics identified in (1) in the context of agent conversation; 4) Appropriate scenarios where conversation with agents could be used. The participant-led characteristics discussed in (1) were written down by the interviewer and discussed again during topic (3). These characteristics and the contexts in which they were discussed were participant generated rather than guided by the interviewer. Furthermore, participants were not prompted to discuss conversations with agents prior to being asked by the interviewer. Following the interviews, participants completed a brief online questionnaire to gather demographics. They were then debriefed and thanked for participating.

The audio recordings were transcribed and analysed in a systematic qualitative interpretation using Inductive Thematic Analysis [9]. Initial coding was conducted by two researchers working independently. We began with an inductive approach, then grouped themes under central topics, informed by the interview guide. Once all data was coded and a set of initial themes was considered, a data session with additional researchers took place. In this session, the researchers closely reviewed the coding and preliminary themes from the data. These themes were further refined by the initial data coders. The transcribed quotes presented in the findings below are representative of the themes they discuss. This follows the same practise in similar qualitative HCI research [20, 35]. All researchers had a background in HCI, with experience in conducting qualitative data analyses. In the following section we present the findings from our analysis.

4 Findings

Purposes of Conversation

Echoing the literature [15], two broad purposes for conversation were clearly shown in the data. These reflected social and transactional goals at the core of most discourse.

Social Purposes The desire to socialise with others was commonly discussed as a principle aim of conversation. Conversation was seen a way of establishing, maintaining and building social bond. Our participants felt conversation was imperative for getting to know people, forming social groups and deepening relationships.

"You can go a bit deeper into knowing people. I would say talking and having conversation is the biggest part of knowing somebody." [P101]

Within social conversation more frivolous talk or 'having the craic' may be dynamically interspersed with more serious topics of discussion.

"You can have very serious conversations...but then you're just talking absolute shite… having the craic and there's so much historical context for that conversation so it's enjoyable." [P106]

Transactional Purposes Juxtaposing the social nature of conversation, participants described times when discourse is more goal-oriented, allowing for people to gather information they need to complete a clearly delineated task or objective. In these, the conversation may shift or end after the speaker feels their goal have been achieved.

"In some conversations you're just trying to elicit information and there's a very clear purpose… There's a very clear short-term objective to the conversation. It's very brief, very goal-oriented. Once you achieve that you just move on." [P109]

Attributes of Conversation

Participants described several key attributes that they value in conversation. Both the purpose of conversation and whom a conversation is with can change the importance of these attributes and the role they play. A number of these attributes and their purposes align with issues discussed in existing linguistics literature [12, 16, 30, 31, 57].

Mutual Understanding & Common Ground Establishing common ground with others was often mentioned as an integral feature of good conversation. Participants stressed the importance of understanding the intent and meaning behind what other speakers are saying beyond. This may involve getting to know others through questioning.

"Sometimes people don't get what you're saying, even though you could be using simple language and talking about simple stuff. It's important that both of you understand what each other is saying, and I don't mean that in terms of like a language barrier." [P105]

"The questions you ask and the questions people ask you even more so are telling about who that person is, what information is it that they're looking for. You're always trying to find a common ground." [P108]

As well as providing a mutually understood focus during interaction, a knowledge of others supports attempts to reach a common understanding.

Trustworthiness Trustworthiness was also discussed as key for conversation through being a key foundation for growing common ground and subsequently sustaining long-term relationships.

"Trustworthy is one of the major characteristics for me. I wouldn't really bother having a long conversation with somebody I knew wasn't trustworthy. Having personal conversations with them would be quite disastrous." [P103]

"You build up this reservoir of knowledge about this person where one day [you] go 'you know what? I can really trust this human being.' When you get to this stage there's desire to keep them in your life." [P106]

Having trust in a partner seems to be a gateway to open the possibility of more personal conversations. Without this trust, such conversations, which are common in long term and deep relationships, may be seen as inappropriate.

Active Listening Active listening was also an important partner attribute in conversation. Participants described that paying attention, demonstrating engagement and a willingness to participate in conversation was important in a two-way interactive dialogue.

"Yeah, we know when they are not following what we are saying so paying attention to somebody who is speaking is very important." [P103]

"Wanting to continue the conversation - wanting to know what comes next and being interested in what the other person is saying. I guess feeling that it's two way as a conversation… a conversation has to go both ways as opposed to one person participating." [P109]

Conversational partners need to demonstrate both good listenership and understanding. In additional to verbal cues, this can be helped by nonverbal feedback like eye contact.

"I definitely prefer when people look me in the eye. I find it very unsettling if you know you're speaking to somebody and they can't make eye contact with you even if it's just fleeting eye contact." [P108]

Humour Humour was also commonly mentioned. Participants proposed that it scaffolds and adds more substance to

discussions. They also remarked how humour can be a key conversational driver.

"I think a bit of humour is fundamental to good conversation." [P112]

"Conversations can just be humorous as well. I find there tends to be some kind of substance underlying it otherwise it's not engaging." [P111]

Although, humour may be an important conversational characteristic, P111's comments also highlight that humour itself needs substance and relevance to the conversation. The comment alludes to humour's ability to soften serious intentions or deliver substantive messages in conversation, in a way that may save face for the speaker.

Building Relationships with Conversation

Participants also discussed the differences in conversational needs depending on the types and stages of relationships they have with interlocutors.

Conversing with Friends When conversing with friends, participants mentioned relying heavily on shared experiences and a history of trust. Because of this trust, our participants regularly conversed with close friends to get advice or alternate perspectives, using them as sounding boards to offload issues and release personal tension.

"I have a friend who I really trust a lot and respect, so I usually make sure I call her and get her opinion about the situation." [P103]

"You feel there is someone for you. If you are in a problem and if you need a second advice, then they know [you] from maybe years together." [P114]

"Sometimes when you're narrating an incident, you're just doing that to get a load off your mind." [P103]

The type of topics covered with close friends were more personal in nature compared to other types of conversational partners (see followin subsection). Such conversations seem built upon a shared repository of memories and experiences, that are co-constructed.

"You have intimate conversations where you would be talking about something personal to them…friend's grandmother was sick last week. Something like that that is intimate, private, discreet." [P109]

"Your conversation leads to memories, building memories together, and that's a huge part of having a friendship that you can reference back and say we did this together and we did that together." [P106]

With very close friends, participants noted that there may not be a need for constant conversation, as silence may symbolise a level of comfort between two friends.

"…with my very close friends, I don't need to talk - it's not like we talk at every given moment.…" [P101]

Conversing with Acquaintances and Strangers When compared to conversations with close friends, those with strangers and casual acquaintances often had a more superficial and functional purpose. Common ground is formed from shared context (e.g. the workplace) or from shared identity with the interlocutor (e.g. colleagues) and used as the anchor for conversational topics and content.

"With acquaintances it's not as personal…it would be more about the common area where we know each other. If we are colleagues, we are probably talking about a meeting we attended or common acquaintances at work." [P103]

"I usually talk about public life, like work but on a very superficial level." [P104]

With acquaintances and strangers, topics covered were more superficial and general in nature, focusing on current events, shared contexts or the need to share information.

"…I suppose we talk…like surface level stuff again like the weather…or maybe sharing a complaint like if you are waiting for a bus" [P108].

The content of these conversations may be more limited compared to friends as participants noted concern as to what topics and behaviour are deemed appropriate.

"…it sometimes has to be more stilted because you can't assume everyone will find the same thing appropriate. You have to minimise the scope of the conversation." [P111]

"…the further you get out from your inner circle [of friends] the constraints come in on what you're going to speak about." [P106].

These interactions may be driven by a perceived social norm to initiate or respond to conversational approaches. For instance, not doing so may seem impolite, particularly if initiated by another speaker. Similarly, participants may use conversation to reduce feelings of awkwardness, especially in periods of silence.

"A lot of people are uncomfortable with silence. When you get very far from the core of very close friends, that discomfort you feel with silence grows… With other people I don't spend a lot of time with, it's very important that the time is spent talking." [P101]

"In terms of day-to-day, you know just meeting people on the street, to me it's social expectation…Even saying nothing is saying something." [P111]

Engaging in conversation, even if just small talk, can help make others and oneself feel comfortable in these situations. Especially with strangers, small talk and transactional dialogue were clearly the main drivers of conversation.

Transition Towards Friendship Participants identified conversation as a fundamental tool used to transition towards friendship, allowing them to 'gauge' one another. Being able to share vulnerabilities and personal information was seen as an important step towards developing mutual trust and bond with others.

"It's [conversation] how you gauge a connection and a forge a pathway where you can understand what is going on with this person. I think bonds can be created with much conversation for sure." [P106]

"I think sharing a vulnerability no matter how small or even acknowledging a sameness… That shows you that the person trusts you if they're willing to tell you something about themselves." [P108]

Developing common ground, through discovering shared interests and traits, was again seen as important to relationship transition. This was seen to develop further through repeated and shared experiences with others.

"Only if we are open about discussing all of these things we figure out what our common interests are and if there are common interests of course then the conversation flows even better and by spending more and more time you become a friend." [P103]

Purposes of Agent-Based Conversation

Transactional over Social Conversations There was a marked difference in the way that participants discussed having conversations with agents compared to conversations with other people. Conversations with agents were almost universally described in functional terms. Their status as an agent meant participants perceived a high barrier to reaching the more social and emotional connections seen in human conversation.

"I would still think of a conversational agent as a tool." [P109]

"Because it's a machine you can't make an emotional connection with it." [P102]

Participants identified that the concept of conversation may need to be reconsidered or defined around different parameters in agent-based interaction. Emulating human interaction was perceived as difficult, if not impossible, with users questioning whether this emulation was even desirable.

"So probably the conversation with a machine should be characterised by different aspects...like for example the clarity of the conversation." [P104]

"What we see as conversation, I don't think we can replicate it so I think there has to be new parameters for what a conversation is with a machine." [P106]

Attributes of Conversation with Agents

When asked to reflect on the attributes mentioned in human conversation, it became clear that these were conceptualised in markedly different terms when applied to human-agent conversation

One Way Understanding and Personalisation over Common Ground Participants described building common ground and mutual understanding as integral to conversation with other people. Yet, there was a clear difference in the role that participants perceived common ground should play in conversational agent interaction. Some participants felt displeased when considering common ground operating in a similar way to human conversation.

"I'd be quite upset if I thought that I had common ground with a computer." [P108]

Instead, common ground was conceptualised as personalisation, where information would be remembered by the agent to tailor their experience. Participants noted that, over time this could create an illusion of common ground between a machine-like agent and user.

"I would find it very difficult to comprehend common ground with a machine. If you were to personalise your machine you might have the perception of common ground. For example, I like rugby so if my machine used rugby analogies explaining things to me I'd perceive it as having a common ground." [P105]

"The more you interact with agents the more they learn about you so the more personalised it becomes...So machines keep learning from what you give as input to them." [P103]

Common ground was not perceived as co-constructed in agent dialogue. Participants felt that the system should lead this process, with the user making little effort to support this.

Functional Trustworthiness and Privacy Trustworthiness in human conversations was linked to sharing personal information and vulnerabilities to increase social bond. In agent conversations, trustworthiness was discussed exclusively in utilitarian terms. Responses related to security, privacy, and transparency over emotional trust.

"Trustworthy when it comes to a machine is more about the security features that are built into it." [P103]

"I think trust definitely in regards to data...you know is this machine recording our conversation? How is that information being used? How are you using my data? Who has access? I think that's where trust might come in." [P108]

The definition of trustworthiness clearly lacked the emphasis on emotional trust seen when discussing human conversational characteristics. That said the conceptualisation

of trust in this context may still act as a gateway to further interaction, in that issues of efficiency, reliability and security may be important for frequent long-term use of conversational agents.

Accurate Listening Again, like other attributes, participants defined the role of the agent as a listener in more functional terms. Participants emphasised the need for the agent to understand them clearly and quickly, ideally without repeating themselves. Many of the comments focused around speech recognition performance.

"I think the voice recognition and not having to repeat yourself would be more of the receptive side of the machine...It's not bettering your experience if you have to sit and repeatedly ask." [P108]

"I suppose that's number one. I want the agent to understand what I'm saying and be able to parse it properly." [P111]

Humour as a Novelty Feature While humour was seen to scaffold human conversations, participants described humour as more of a novelty feature that can help make interactions with agents more interesting. Agents were seen to lack the 'organic' process of humour seen between people.

"It [humour] makes the interaction interesting and you want to keep using these kind of devices which makes it fun for you to use it." [P103]

"I don't think you're going to get Humour. I think it's hit or miss. I know you can ask Siri to tell you jokes and sometimes they do and it's usually on the dad jokes level. It's not actually funny the way an organic conversation would be between two humans." [P117]

While the novelty of humour with agents was noted as a positive feature it was rarely described as necessary in the same way it was for human conversation.

Relationship Building in Agent-Based Conversation

Becoming Friends with Conversational Agents As presented earlier, our analysis showed that current human-agent conversations serve primarily transactional purposes. From a relationship point of view, this echoes the transactional interactions people have with strangers or casual acquaintances. Conversation was identified as the cornerstone of becoming close friends in human interaction. A key question arose as to whether and how this can be accomplished with agents. In our data, participants could not overcome the functional purposes of human agent conversation and were resistant to the idea of becoming friends with something inherently machine-like. Conversational agents were consistently considered as tools and assistants available to serve and accommodate people. Motivations for building a different relationship with them were questioned.

"I don't know why I would want my tool to be vulnerable, or be intimate, I think like you said it has– like if I wanna put a nail in a table, I get a hammer. If I want to find out how to get to a particular place you'd put it into, Siri or something like that." [P109]

"I dunno if you really want to sit and have a conversation that didn't require something out of the machine like the weather or turn this on or make an appointment for me or that kind of thing but... you're not gonna make friends with a machine so…" [P117]

Participants dismissed friendships with machines as 'un-normal' or were reluctant to envisage having conversations with them in the same way as they would with their friends.

"I mean I don't enjoy communicating with machines...when machines are like trying to become human that's just not ok...because. If you're talking to a machine to get their perspective on something, it'd be bit, un-normal." [P112]

"It's not like you can chat with a chatbot about how you feel and why your morning sucked." [P101]

Conversing with agents was perceived as innately different to interacting with people, meaning friendship building could not be accommodated

"But it's still not the same as with a human 'cause the machine is gonna be like...running through algorithms to go what are we talking about– keep in this range so they're not gonna interrupt halfway through go oh my god I forgot to tell you about this, or before I forget I have to tell you about that, that kind– a machine is not gonna do that." [P117]

Relationship Tensions Participants pointed to potential conflicts that can arise if people were to be friends with machines. Building a relationship with a conversational agent was seen to require significant time and effort together with a potential shift of the nature of friendship.

"… we're going to have to really reconfigure how we think about what friendship is for us for this to work…If that's what people would want… because it's not like having a pet…It will be very very interesting to see if people do start adopting… machines as companions in the home." [P106]

The prevailing perception of a master-servant relationship invokes key question in terms of how to reconcile ordering around a friend, or what would occur if an agent was to decline your request?

"It's a strange thing, on one hand you're sort of becoming a friend, but on the other you can order it to do what you want …or ask it what you want. In that sense it would be difficult to ever really think of it as a close acquaintance. Then it would have the right to say no to you. And then people throw those things out the window" [P113]

A further tension involved the monetary value of agents for companies and concerns around how monetary incentives will operate in the context of a human-agent friendship.

"It depends, to my idea, the main reason behind chatbots, is pretty much money - there is not really a conversation going on. It's more people want to create chatbots because they want to have a more flexible customer care - maybe less expensive. It's a good application of a nice and powerful technology but it's different from conversations itself, right?" [P101]

Potential Scenarios for a Human-Agent Relationship Despite initial reluctance participants did see opportunities for a human-agent relationship to be of value. Such scenarios involved people who are isolated such as elderly or struggling with mental health issues and could benefit from conversation.

"Let's say for example it's a case... of anxiety. And you're going, I have to go to this work thing, but I don't want to go. And the machine will start telling you what the benefits are of going. Maybe it's better for your career, maybe you'll, get more chance... you'll be seen. All this type of stuff." [P113]

"I think that those are the only two, like task based and keeping the person company." [P112]

That said, when considering potential scenarios where conversational agent could be valuable, participants primarily described functional applications such as controlling home appliances, monitoring health, scheduling appointments or dealing with daily personal administrative tasks.

"If you have something in the home that is able to guide and navigate people through their illness and teach them how to self-manage better. If it's as simple as reminding them to take their medications... telling them about hospital appointments" [P104]

"Not nagging you but at least giving you prompts and giving you options of how you might actually conveniently do this...These things should definitely be linked into all that. They're aware of your taxes as well holy God, nobody understands about taxes like. It's all stuff you have to find out, that you know there's this machine that's able to take all of that admin stuff and even just organise it for you." [P106]

5   Discussion

Our study aimed to identify important conversational characteristics in human conversation and how these may vary when applied in conversations with agents. Conversational properties in human-human dialogue are well known and the current uses of speech agents are also well mapped. Yet people's perceptions of conversing with machines (not just IPAs) are less understood. Through asking participants to reflect on the properties of conversation in both human and agent contexts, we aim to contrast the nature of human-human and human-agent conversation. Our findings show that compared to the perceived social and transactional purposes of human-human conversation, agent-based conversation was conceptualised in almost purely transactional terms. Echoing seminal work in linguistics [12, 16, 30, 31, 57], important characteristics of human-human conversation focused on mutual understanding and common ground, trust, active listenership, and humour. When reflecting on these characteristics in agent contexts people described them in highly functional terms. In conversations with agents, common ground development was viewed as a one-way process related to personalisation. Although this brings an illusion of mutual shared knowledge integral to common ground [16, 17], participants clearly do not see this as being collaboratively built, negotiated and updated during agent dialogue, juxtaposing grounding processes in human conversation [18]. Trust and listenership were defined in terms of system performance rather than important precursors for social bond, whilst humour was seen as a novelty rather than an integral component of conversation.

Conversation was integral to developing long-term friendships between people. Conversely, participants fundamentally questioned the desire and ability to befriend or converse with an agent in the same way. They identified key conflicts in trying to do so (e.g. issues with master-servant dynamic, dissonance between monetary incentives and agent-friendship development). Core to many issues was the status asymmetry seen between users and conversational agents, where the agent is perceived as a tool or servant for the user. This creates a fundamental barrier to developing long-term relationships. That said, there were limited contexts where participants felt that conversation with agents would be beneficial, such as to support people who are isolated or struggling with mental health issues.

Current Perceptions as a Barrier to Conversational Agent Interaction

Our data emphasised a fundamental mismatch in perceived status between agent and user where the agent was considered a user-controlled tool rather than a potential companion or social equal. This, in tandem with perceptions of functional ability [39] may restrict the types of conversations that users perceive as appropriate or possible to have with an agent at present [19]. Critically, our findings suggest that social aspects of conversational interaction are currently absent from people's perceptions of what conversational agents can and should be capable of performing. These expectations support the notion that an agent is a basic dialogue partner [8] lacking in humanlike conversational capabilities [20, 35, 42]. Our participants personally expressed no desire to build bonds with conversational agents (although they

expressed the view that other user group may find conversational features useful). The lack of enthusiasm for bonding may stem from the core belief that agents are poor dialogue partners that should be subservient to the user. It may also lie in the perception that there is no support for social dialogue in the current infrastructure for conversation [22], as more social and conversational dialogue is currently lacking in current VUIs [42]. This type of perception may be anchored by current experiences of IPAs, which are not designed to satisfy interpersonal goals, fuelling stereotypes around conversational agent abilities. The basic view of agent capability may be one of the most significant current barriers to users embracing or utilising conversational agents for social goals.

Reframing the Concept of Conversation in Agent Interaction

Our findings support the view that the paradigm of conversation with agents needs to be reframed [42, 45]. It is clear that participants in this paper categorised conversation with agents as almost exclusively task-oriented and transactional, echoing findings in other literature [20, 35, 42]. Current commercial IPAs are not necessarily designed to deliver social conversation. However, there remain contexts where, because of the need to foster a long term and personal human-agent relationship, conversational agents may need orienting towards addressing interpersonal and social goals. Yet this may not need to emulate the form or outcomes of human conversation. Rather than simulating human conversational abilities in the hope of successful social conversation with users, our findings suggest that we need to treat human-agent interaction as a new genre of conversation, with its own rules, norms and expectations. As articulated in our data these may be more functional and utilitarian in nature, with little emphasis on the relational growth or emotional outcomes seen in human conversation. Our data suggests that the conversational content and structure present in interactions we have with strangers or acquaintances may be a good starting metaphor for social agent conversations.

Indeed agent conversations may mirror more limited service-oriented or 'front desk' encounters between people. Talk is sequential [48], and in limited encounters, conversation is often sequentially organised around adjacency pairs signifying requests for services followed by provisions of services [49]. They can also include openings to signifying the beginning of interactions (similar to IPA 'wake words') and closings to signify interactions ending. Human-human service encounters may also include an optional interrogative series, where more information is gathered [32]. These features differ from modern speech technology which are often limited to isolated adjacency pairs without closings [42]. However, these conversational norms and indeed expectations of speech technology will likely be dynamic, shifting as long-term and multiple-turn agent use become more commonplace in contexts where they are designed to address social needs and not simply fulfill service requests. These conversational norms and expectations will likely be dynamic, shifting as long-term agent use becomes more commonplace in contexts where they are designed to address social needs.

The importance of context in shaping the content and norms of agent conversation cannot be overstated. For instance, what people deem appropriate to divulge or discuss conversationally with agents may differ markedly in private and public settings. Current users are unlikely to engage with IPAs in public [20, 35], noting social embarrassment and awkwardness [1, 20]. Users have fewer concerns divulging private information when using VUIs in private compared to social spaces [1, 24]. We identify a clear distinction between human and agent based conversation in terms of its perceived norms, rules and expectations. The context of interaction is no doubt likely to impact these and further work should look to explore this impact.

Agent Conversation for Social Goals

Our work questions the extent to which imbuing conversational attributes in agents will lead to similar benefits and relationships to those in human communication. However, in some contexts conversational capabilities may help facilitate interaction and use. For example, research in social robotics [47] shows that elderly users may benefit from social capabilities in a system, leading them to disclose personal stories. A sensitivity to the context of interaction, an understanding of the type of conversation required, and the purposes behind the interaction are integral for this to succeed. As suggested in our data, conversation with agents may not be the best approach to use to develop bond. This could be achieved through exploiting other modalities, especially in more embodied systems (e.g.[30]). Through gesture and facial expressions cues (e.g. [10, 13, 33]) embodiment may even help support speech-based conversation.

Limitations and Future Work

This paper identifies common characteristics people view as important in conversation and how these differ when applied to conversations with agents. Participants saw interpersonal relationships as immaterial when engaging with conversational agents. These views may come from a lack of familiarity with agents that are designed to engage in conversation. In our data it is clear that participant views were grounded by existing IPA experiences and the types of interactions those agents facilitate (e.g. short-term, sporadic and transactional). Future research should explore ways to get users to focus on envisioning longer, sequential interactions where conversational capabilities are more sophisticated [44, 46]. Although

we did not get participants to interact with conversational agents during the research, getting users to engage with a (real or simulated) conversational agent, may act as a useful trigger to support generation of characteristics and scenarios. These approaches could be particularly valuable for the design of future conversational agents, identifying issues that may arise in their implementation. Future research would also benefit from recruiting participants who are less familiar with technology. Most participants in this study described themselves as intermediate, advanced or expert users of technology. People who are further away from understanding how IPAs and similar technology works may have different opinions towards agent-based conversations. Recent studies have indicated that certain users groups tend to anthropomorphise their devices, particularly older adults and children (e.g.[43, 56]). Users such as these may perceive conversation with machines differently to the sample population in this paper.

6 Conclusion

Conversations play an essential role in building and maintaining relationships with other people. However, this is not seen as important or even desirable in most current scenarios with conversational agents. While many concepts of good human-human conversation are discussed as important in human-agent conversations, how they are described is markedly different. Participants describe mutual understanding and common ground, trust, active listenership, and humour as crucial social features in human conversations. In agent conversations, these are described almost exclusively in transactional and utilitarian terms. Our findings suggest there may be a limit to the extent interactions with agents can mirror those with other people. These seem to be influenced in particular by existing agents whose interactions stray from common definitions of conversation. However, there may be specific application areas where conversation may be appropriate if not essential between humans and agents, particularly in areas such as healthcare and wellbeing, where the nuances of contexts and demographics need to be considered. This paper suggests that conversational agents can be inspired by human-human conversations but do not necessarily need to mimic it. Instead, human-agent conversations may need to be seen as a new genre of interaction.

References

[1] Ali Abdolrahmani, Ravi Kuber, and Stacy M Branham. 2018. Siri Talks at You: An Empirical Investigation of Voice-Activated Personal Assistant (VAPA) Usage by Individuals Who Are Blind. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 249–258.

[2] Frank Bentley, Chris Luvogt, Max Silverman, Rushani Wirasinghe, Brooke White, and Danielle Lottrjdge. 2018. Understanding the Long-Term Use of Smart Speaker Assistants. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 91.

[3] Timothy Bickmore and Justine Cassell. 2001. Relational agents: a model and implementation of building user trust. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 396–403.

[4] Timothy Bickmore and Justine Cassell. 2005. Social dialngue with embodied conversational agents. In *Advances in natural multimodal dialogue systems*. Springer, 23–54.

[5] Timothy W Bickmore, Lisa Caruso, and Kerri Clough-Gorr. 2005. Acceptance and usability of a relational agent interface by urban older adults. In *CHI'05 extended abstracts on Human factors in computing systems*. ACM, 1212–1215.

[6] Timothy W Bickmore, Suzanne E Mitchell, Brian W Jack, Michael K Paasche-Orlow, Laura M Pfeifer, and Julie O'Donnell. 2010. Response to a relational agent by hospital patients with depressive symptoms. *Interacting with computers* 22, 4 (2010), 289–298.

[7] Timothy W Bickmore and Rosalind W Picard. 2004. Towards caring machines. In *CHI'04 extended abstracts on Human factors in computing systems*. ACM, 1489–1492.

[8] Holly P Branigan, Martin J Pickering, Jamie Pearson, Janet F McLean, and Ash Brown. 2011. The role of beliefs in lexical alignment: Evidence from dialogs with humans and computers. *Cognition* 121, 1 (2011), 41–57.

[9] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[10] Cynthia Breazeal. 2003. Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies* 59, 1 (July 2003), 119–155. https://doi.org/10.1016/S1071-5819(03)00018-1

[11] Gillian Brown and George Yule. 1983. *Discourse Analysis*. Cambridge University Press. Google-Books-ID: 9gUrbzov9x0C.

[12] Penelope Brown and Stephen C Levinson. 1987. *Politeness: Some universals in language usage*. Vol. 4. Cambridge university press.

[13] A. Bruce, I. Nourbakhsh, and R. Simmons. 2002. The role of expressiveness and attention in human-robot interaction. In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, Vol. 4. 4138–4142 vol.4. https://doi.org/10.1109/ROBOT.2002.1014396

[14] BusinessWire. 2018. Global Intelligent Virtual Assistant Market 2018-2023:. https://www.businesswire.com/news/home/20180723005506/en/Global-Intelligent-Virtual-Assistant-Market-2018-2023-Market

[15] C. Cheepen. 1988. *The predictability of informal conversation*. Pinter. https://books.google.ie/books?id=ai23AAAAIAAJ

[16] H.H. Clark. 1996. *Using Language*. Cambridge University Press. https://books.google.ie/books?id=DiWBGOP-YnoC

[17] Herbert H Clark. 1992. *Arenas of language use*. University of Chicago Press.

[18] Herbert H Clark, Susan E Brennan, et al. 1991. Grounding in communication. *Perspectives on socially shared cognition* 13, 1991 (1991), 127–149.

[19] Leigh Clark. 2018. Social Boundaries of Appropriate Speech in HCI: A Politeness Perspective. In *Proceedings of the 32nd British Human Computer Interaction Conference*. British Computer Society. https://doi.org/10.14236/ewic/HCI2018.76

[20] Benjamin R Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira.

2017. What Can I Help You With?: Infrequent Users' Experiences of Intelligent Personal Assistants. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, 43.

[21] L. Devillers, S. Rosset, G. Dubuisson Duplessis, L. Bechade, Y. Yemez, B. B. Turker, M. Sezgin, E. Erzin, K. El Haddad, S. Dupont, P. Deleglise, Y. Esteve, C. Lailler, E. Gilmartin, and N. Campbell. 2018. Multifaceted Engagement in Social Interaction with a Machine: The JOKER Project. In *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*. 697–701. https://doi.org/10.1109/FG.2018.00110

[22] Hugh Dubberly and Paul Pangaro. 2009. What is conversation, and how can we design for it? *interactions* 16, 4 (2009), 22–28.

[23] Robin Dunbar. 1998. *Grooming, gossip, and the evolution of language*. Harvard University Press.

[24] Aarthi Easwara Moorthy and Kim-Phuong L Vu. 2015. Privacy concerns for use of voice activated personal assistant in the public space. *International Journal of Human-Computer Interaction* 31, 4 (2015), 307–335.

[25] Suzanne Eggins and Diana Slade. 2005. *Analysing Casual Conversation*. Equinox Publishing Ltd. Google-Books-ID: cTEkm6WzfNgC.

[26] Hao Fang, Hao Cheng, Elizabeth Clark, Ariel Holtzman, Maarten Sap, Mari Ostendorf, Yejin Choi, and Noah A Smith. 2017. Sounding Board – University of Washington's Alexa Prize Submission. (2017), 12.

[27] Emer Gilmartin, Marine Collery, Ketong Su, Yuyun Huang, Christy Elias, Benjamin R. Cowan, and Nick Campbell. 2017. Social Talk: Making Conversation with People and Machine. In *Proceedings of the 1st ACM SIGCHI International Workshop on Investigating Social Interactions with Artificial Agents (ISIAA 2017)*. ACM, New York, NY, USA, 31–32. https://doi.org/10.1145/3139491.3139494

[28] Emer Gilmartin, Benjamin R. Cowan, Carl Vogel, and Nick Campbell. 2018. Explorations in multiparty casual social talk and its relevance for social human machine dialogue. *Journal on Multimodal User Interfaces* (Sept. 2018). https://doi.org/10.1007/s12193-018-0274-2

[29] Rachel Gockley, Allison Bruce, Jodi Forlizzi, Marek Michalowski, Anne Mundell, Stephanie Rosenthal, Brennan Sellner, Reid Simmons, Kevin Snipes, Alan C Schultz, and others. 2005. Designing robots for long-term social interaction. In *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*. IEEE, 1338–1343.

[30] Erving Goffman. 1967. *Interaction Ritual : Essays in Face-to-Face Behavior*. Routledge. https://doi.org/10.4324/9780203788387

[31] Adam Jaworski. 2014. Silence and small talk. In *Small talk*. Routledge, 130–152.

[32] Mardi Kidwell. 2000. Common ground in cross-cultural communication: Sequential and institutional contexts in front desk service encounters. *Issues in Applied Linguistics* 11, 1 (2000).

[33] Yoshinori Kuno, Kazuhisa Sadazuka, Michie Kawashima, Keiichi Yamazaki, Akiko Yamazaki, and Hideaki Kuzuoka. 2007. Museum Guide Robot Based on Sociological Interaction Analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, New York, NY, USA, 1191–1194. https://doi.org/10.1145/1240624.1240804

[34] Rosemarijn Looije, Mark A. Neerincx, and Fokie Cnossen. 2010. Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. *International Journal of Human-Computer Studies* 68, 6 (June 2010), 386–397. https://doi.org/10.1016/j.ijhcs.2009.08.007

[35] Ewa Luger and Abigail Sellen. 2016. Like having a really bad PA: the gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 5286–5297.

[36] Bronislaw Malinowski. 1994. The problem of meaning in primitive languages. *Language and literacy in social practice: A reader* (1994), 1–10.

[37] Michael McTear, Zoraida Callejas, and David Griol. 2016. *The conversational interface: Talking to smart devices*. Springer.

[38] Rishabh Mehrotra, A Hassan Awadallah, AE Kholy, and Imed Zitouni. 2017. Hey Cortana! Exploring the use cases of a Desktop based Digital Assistant. In *SIGIR 1st International Workshop on Conversational Approaches to Information Retrieval (CAIR'17)*, Vol. 4.

[39] Roger K. Moore. 2017. Is Spoken Language All-or-Nothing? Implications for Future Speech-Based Human-Machine Interaction. In *Dialogues with Social Robots: Enablements, Analyses, and Evaluation*, Kristiina Jokinen and Graham Wilcock (Eds.). Springer Singapore, Singapore, 281–291. https://doi.org/10.1007/978-981-10-2585-3_22

[40] Cosmin Munteanu, Pourang Irani, Sharon Oviatt, Matthew Aylett, Gerald Penn, Shimei Pan, Nikhil Sharma, Frank Rudzicz, Randy Gomez, Ben Cowan, et al. 2017. Designing Speech, Acoustic and Multimodal Interactions. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 601–608.

[41] Ioannis Papaioannou, Amanda Cercas Curry, Jose L Part, Igor Shalyminov, Xinnuo Xu, Yanchao Yu, Ondrej Dušek, Verena Rieser, and Oliver Lemon. 2017. Alana: Social dialogue using an ensemble model and a ranker trained on user feedback. *Alexa Prize Proceedings* (2017).

[42] Martin Porcheron, Joel E Fischer, Stuart Reeves, and Sarah Sharples. 2018. Voice Interfaces in Everyday Life. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 640.

[43] Amanda Purington, Jessie G Taft, Shruti Sannon, Natalya N Bazarova, and Samuel Hardman Taylor. 2017. Alexa is my new BFF: social roles, user satisfaction, and personification of the amazon echo. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2853–2859.

[44] Stuart Reeves. 2012. Envisioning Ubiquitous Computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 1573–1582. https://doi.org/10.1145/2207676.2208278

[45] Stuart Reeves. 2017. Some conversational challenges of talking with machines.. In *In: Talking with Conversational Agents in Collaborative Action Workshop*.

[46] Tom A. Rodden, Joel E. Fischer, Nadia Pantidi, Khaled Bachour, and Stuart Moran. 2013. At home with agents: exploring attitudes towards future smart energy infrastructures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. ACM Press, Paris, France, 1173. https://doi.org/10.1145/2470654.2466152

[47] A. M. Sabelli, T. Kanda, and N. Hagita. 2011. A conversational robot in an elderly care center: An ethnographic study. In *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 37–44. https://doi.org/10.1145/1957656.1957669

[48] Harvey Sacks. 1987. On the preferences for agreement and contiguity in sequences in conversation. *Talk and social organisation* (1987), 54–69.

[49] Harvey Sacks, Emanuel A Schegloff, and Gail Jefferson. 1978. A simplest systematics for the organization of turn taking for conversation. In *Studies in the organization of conversational interaction*. Elsevier, 7–55.

[50] Martin Saerbeck, Tom Schut, Christoph Bartneck, and Maddy D Janse. 2010. Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1613–1622.

[51] Klaus P. Schneider. 1988. *Small talk: Analyzing phatic discourse*. Vol. 1. Hitzeroth.

[52] Brendan Spillane, Emer Gilmartin, Christian Saam, Ketong Su, Benjamin R Cowan, Séamus Lawless, and Vincent Wade. 2017. Introducing ADELE: a personalized intelligent companion. In *Proceedings of the 1st ACM SIGCHI International Workshop on Investigating Social Interactions*

with Artificial Agents. ACM, 43–44.

[53] Adriana Tapus, Maja J Mataric, and Brian Scassellati. 2007. Socially assistive robotics [grand challenges of robotics]. *IEEE Robotics & Automation Magazine* 14, 1 (2007), 35–42.

[54] Janice Y. Tsai, Tawfiq Ammaria, Abraham Wallin, and Jofish Kaye. 2018. Alexa, play some music: Categorization of Alexa Commands. In *Workshop on Voice-based Conversational UX Studies and Design (CHI'18)*.

[55] Laura Pfeifer Vardoulakis, Lazlo Ring, Barbara Barry, Candace L Sidner, and Timothy Bickmore. 2012. Designing relational agents as long term

social companions for older adults. In *International Conference on Intelligent Virtual Agents.* Springer, 289–302.

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

[57] Avner Ziv. 2010. The Social Function of Humor in Interpersonal Relationships. *Society* 47, 1 (Jan. 2010), 11–18. https://doi.org/10.1007/s12115-009-9283-9