

Mapping Perceptions of Humanness in Speech-Based Intelligent Personal Assistant Interaction

Philip R. Doyle

University College Dublin
philip.doyle1@ucdconnect.ie

Justin Edwards

University College Dublin
justin.edwards@ucdconnect.ie

Odile Dumbleton

University College Dublin
odile.dumbleton@ucdconnect.ie

Leigh Clark

University College Dublin
leigh.clark@ucd.ie

Benjamin R. Cowan

University College Dublin
benjamin.cowan@ucd.ie

Abstract

Humanness is core to speech interface design. Yet little is known about how users conceptualise perceptions of humanness and how people define their interaction with speech interfaces through this. To map these perceptions $n=21$ participants held dialogues with a human and two speech interface based intelligent personal assistants, and then reflected and compared their experiences using the repertory grid technique. Analysis of the constructs show that perceptions of humanness are multidimensional, focusing on eight key themes: *partner knowledge set, interpersonal connection, linguistic content, partner performance and capabilities, conversational interaction, partner identity and role, vocal qualities* and *behavioral affordances*. Through these themes, it is clear that users define the capabilities of speech interfaces differently to humans, seeing them as more formal, fact based, impersonal and less authentic. Based on the findings, we discuss how the themes help to scaffold, categorise and target research and design efforts, considering the appropriateness of emulating humanness.

CCS Concepts

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

Keywords

speech interface; voice user interface; intelligent personal assistants; repertory grid; mental models; humanness.

ACM Reference Format:

Philip R. Doyle, Justin Edwards, Odile Dumbleton, Leigh Clark, and Benjamin R. Cowan. 2019. Mapping Perceptions of Humanness in Speech-Based Intelligent Personal Assistant Interaction. In *MobileHCI 2019: 21st International Conference on Human-Computer Interaction with Mobile Devices and Services, October 01–04, 2019, Taiwan, Taipei*. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3338286.3340116>

1 Introduction

Speech interfaces like Amazon Alexa, Google Assistant and Apple Siri are now commonly used through mobile and smart speaker devices. These intelligent personal assistants (IPAs) use speech as the primary means of interaction, lacking embodiment compared to other speech systems such as robots or embodied conversational agents (ECAs). HCI research on these speech interface based intelligent personal assistants (IPAs) has looked to develop understanding of users' expectations, perceptions and experiences. This work suggests that users tend to use human dialogue as an initial interaction metaphor, with expectations of system competence and function being set by the human-likeness of their design [15, 17, 38]. Emulating aspects of humanness in speech interfaces is common [8, 10, 16, 37], although the wisdom of this design choice has been questioned [10, 12, 17, 40] as it may create unrealistic expectations of system capability [38]. That said, currently it is not clear how people conceptualise the humanness of speech-based IPAs and how exactly users perceive that humans and speech interfaces vary when considering humanness as the metaphor for interaction. This knowledge is critical to informing the growing debate around humanness. By specifying the dimensions that users find important in this concept we can identify what type of design decisions may influence user expectations. Comparing humans and speech-based IPAs can also lead to a deeper understanding

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MobileHCI 2019, October 01–04, 2019, Taipei, Taiwan

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6825-4/19/10...\$15.00

<https://doi.org/10.1145/3338286.3340116>

of the gulf between expectation and reality [36, 38] when humanness is used as a heuristic for interaction.

The work presented contributes a user-centered insight into the dimensions that are important to users when considering humanness of speech-based IPAs and identifies how these vary from human dialogue. We do this by conducting a study using the Repertory Grid Technique (RGT) where, following interactions with three dialogue partners (a human, Amazon Alexa and Apple Siri), participants generated constructs that best described those interactions. Through content analysis of participant-generated constructs we identify eight salient dimensions along which humanness is most commonly conceptualised. These focus on *partner knowledge set*, the qualities associated with *interpersonal connections*, the *linguistic content* of the partner's utterances, a partner's *performance and capabilities*, their *identity and role* in the dialogue, the *vocal qualities* they portray and *behavioral affordances* toward a partner. We also outline the key differences between perceptions of humans and speech-based IPAs on these dimensions, with interview data showing users are highly aware of the limitations of IPAs in comparison and that they form a different category of communicative interaction along these dimensions. Our findings are the first to outline the multi-faceted nature of user perceptions of humanness, identifying the key dimensions that need to be considered when designing and researching humanness in speech interface interaction.

2 Related Work

Humanness as an interaction metaphor

Previous work on humanness has generally focused on embodied agents and robots, discussing concepts such as the uncanny valley [41]. Studies have also explored categorizations of human-nature and human-uniqueness [27, 28] and how imbuing systems with personality traits increases perceived human likeness [25, 52]. Work also strongly emphasises how embodied factors mediate perceptions of humanness, showing influence from kinematics [49], agent appearance and embodiment [5, 29, 46, 46], facial expressions and gesture [7, 34]. Current, widely used IPA devices are not embodied beyond the device casing and lights that subtly change colour (i.e., they have no appendages, head or face, nor movement like a robot or avatar), so perceptions of humanness may not be driven by similar considerations. Yet as an overarching design principle for non-embodied speech-based IPAs, humanness is a fundamental concept in understanding interactions with current speech interfaces. In particular, speech system development focuses heavily on emulating human aspects of speech through expressiveness and human-like synthesis [1] or human-based conversational rules and structures [21, 26]. This humanness seems

to give users expectations as to the type of capabilities that a system may have as a dialogue partner [36, 38]. Indeed, incongruence between expectations derived from humanness cues and the realities of interacting with speech-based systems is detrimental to speech interface user experience [17, 38, 40, 45]. Qualitative work investigating the user experience of speech-based IPAs for both power users [38] and infrequent users [17], found that they seem to initially use human dialogue as a metaphor for interaction. The humanness of current speech-based IPA design also leads to inaccurate expectations of system capabilities, which are adjusted after experiencing interaction [36, 38]. This impact of humanness is echoed by Moore et al. [40] who highlights that human-like voices in particular may create overestimations of a system's capabilities, resulting in potential communication breakdown and unsuccessful engagement with systems. Although the humanness communicated by the design of speech interfaces may act as an anchor for perceptions, people do seem to make a fundamental distinction between the capabilities of automated and human dialogue partners. They see machines as more basic or functional in their conversational capabilities [4, 13], acting as "at risk" listeners in conversations [42].

Understanding user's perception towards speech interfaces may be important to unlocking reasons for people's behaviours in interaction. Poor perceptions of system capability are assumed to cause us to adapt our language choices in interaction [2, 6, 35, 39]. They drive us to use fewer fillers and coherence markers [2], more basic lexical choices, and shorter utterances compared to when we communicate with a human [33], although the ubiquity of this role is debated [14, 16].

Understanding perceptions using repertory grids

Understanding what factors influence these perceptions could be critical to understanding fundamental mechanisms in speech interface interaction. The RGT is a research tool designed to facilitate the discovery of what may drive users perceptions towards particular objects and situations. Developed as a tool to support personal construct theory in psychology, the technique allows participants to generate labels (termed *personal constructs*) to describe, conceptualise and compare particular objects of study (termed *elements*) [32]. Participants are exposed to three elements at a time, through a paradigm known as *triading*, where constructs are generated to differentiate two similar elements from a distinct third element. RGT therefore allows researchers an insight into an individual's reasoning and conceptualising process as they work toward understanding their experiences [24]. Historically, it has been used in educational psychology [47] and information design [30]. HCI-related studies have also utilized the RGT to examine perceptions of

website usability [50], perceptions of strategic information systems [9] and perceptions of mobile technologies [22].

3 Research Aims

Humanness plays a critical role in the development of perceptions towards IPAs. Yet concepts of humanness are predominantly discussed in embodied interactions, as a unidimensional concept, with little known about how users conceptualise humanness in non-embodied speech interfaces. Insight into user-generated concepts of humanness is crucial to add nuance to the debate around its role in user experience. Mapping the dimensions that users find important when conceptualising the humanness of speech-based IPAs through a bottom-up construct elicitation paradigm will allow us to identify the design dimensions that may need to be prioritised to ensure accurate user perceptions are generated. Through using RGT, this study aims to: (1) map the dimensionality of the concept of humanness in interactions with speech-based IPAs; and (2) identify the nuances between human and IPA interaction within these dimensions.

4 Method

Using random sampling, twenty-four participants from a European university were recruited via an email circulated across the university. Recruitment posters were also displayed around campus. Each participant was given a €10 honorarium for taking part. Three participants were omitted from the data due to them having difficulty in completing the grids unassisted. Of the remaining twenty one participants (F=9, M=11; Mean age = 23.1yrs; SD=5.49) all were native or near native English speakers. 23.8% (N=5) of the sample reported using speech-based IPAs daily to a few times per week with 14.3% of the sample (N=3) reporting only using speech-based IPAs a few times a month. 38.1% of the sample reported using them rarely (N=8) with 23.8% (N=5) having never used them before the study. Among those that had used speech-based IPAs, Siri (50%) was most commonly used, followed by Google Assistant (31.3%) and Amazon Alexa (18.8%).

Procedure

Upon arriving at the lab, all participants were fully briefed about the nature of the study and were asked to provide consent to take part. After completing demographics, participants were then given instructions as to the purpose of the study. Participants were informed that during the session they would have a brief interaction with three dialogue partners where they would ask each of them in turn a set of pre-determined questions (Familiarisation Phase). They were then told that after their interactions they would take

part in a semi-structured interview where they would identify a series of constructs relevant to these interactions (Construct Elicitation Phase). After this, participants were asked to rate each interaction in relation to these constructs on a sliding scale (Rating Phase). The phases are described in detail below.

Familiarisation Phase Each participant completed a familiarisation session whereby they interacted with three elements. The elements selected included Siri (accessed on an iPhone SE smartphone), Alexa (accessed through an Echo Dot device), and a human (a member of the research team). The speech-based IPA elements were chosen because of their popularity and how they varied in their use of multimodality in interaction. Alexa, via the Echo Dot, predominantly uses speech, with minimal visual output through using a single coloured ring on the top of the device. Siri primarily uses speech as a primary interaction modality but also uses significant visual feedback through the smartphone's touchscreen. The human was included to act as a comparator and ensure that constructs would relate to humanness. For ecological validity, human and speech-based IPA responses were unconstrained. Familiarization was completed with all three elements prior to the construct elicitation phase. The order in which elements were addressed was randomised between participants to control for recency and order effects. Only minimal introductions between the human and participants were made by the lead researcher to confine the interaction as much as possible within the predefined questions.

To facilitate interaction with each element, participants were tasked with asking each dialogue partner nine pre-set questions. These were devised by the authors to emphasise differences in the way humans and speech-based IPAs can communicate, so as to ensure a wide set of comparisons could be made. The questions used are shown in Table 1. The order in which elements were addressed was randomised between participants.

Construct Elicitation Phase Next, through a semi-structured interview, participants were asked to compile a list of *implicit constructs* that they felt best described the key similarities and differences between their interactions. These are critical constructs that inform the participant's own understanding of their experiences [24, 32]. Participants were asked to focus on the communicative abilities and qualities of each element. They were also asked to provide context and reasoning around why they were choosing certain words and how it related to their interactions. After an exhaustive list had been compiled (averaging 11-13 constructs), participants were asked to devise a dichotomous *emergent construct* for each *implicit construct* elicited. Participants were informed that this did not have to be a direct

Table 1: Question types with examples

Question/request type	Question/request format
Conversational	<p><i>How are you today?</i></p> <p><i>Where are you from?</i></p> <p><i>Tell me a joke</i></p>
Information retrieval	<p><i>Who is [insert famous person's name]?</i></p> <p><i>What is the square root of [insert three digit number]?</i></p> <p><i>How do I get to the City Centre from here?</i></p>
Subjective/opinion-based	<p><i>Do you like [insert favorite genre of music]?</i></p> <p><i>Can you recommend a place to eat [insert favorite food when eating out]?</i></p> <p><i>What do you think of [insert famous person's name - same as before]?</i></p>

antonym - although it could be if appropriate - but should be based on the context in which they meant the original word. From this a grid of implicit-emergent construct pairs were gathered for each participant.

Rating Phase In this final stage, participants were asked to rate where they felt each element fell between each of the implicit-emergent poles (construct pairs) elicited. To do this, participants were presented with a grid containing all constructs produced during the construct elicitation phase. Implicit constructs were entered into the left hand column, emergent constructs into the right hand column, and both columns were connected with a straight line, 134mm long. Participants were asked to place a mark on the line where they felt each element sat on a spectrum between the construct poles. This data was used to help interpret whether the implicit or emergent construct was more closely associated with human or IPA based experiences. They were also used to provide additional context and help interpret construct meanings, further supporting the categorisation process.

All sessions were recorded and transcribed. Upon completing the session, users were debriefed as to the aims of the study and were thanked for taking part.

Analysis

A bootstrapped content analysis of the construct pairs [31] was conducted to identify consistent themes among the constructs generated. This categorisation helps identify overarching meanings or characteristics across the constructs [31]. Following guidelines by [31] the analysis was conducted over three phases.

Phase 1 Three of the authors with experience in qualitative and/or speech interface research independently assigned initial thematic tags to each implicit-emergent construct pair. This was based on semantic assessment of construct labels, using participant explanations from the interview

data and ratings to ensure the participants' intended meaning was captured. From this phase, an initial understanding and conceptualisation of common themes present in the constructs was produced.

Phase 2 The themes from Phase 1 were then discussed by two of the annotators from the initial coding phase and two speech interface researchers who had not been involved in the initial coding, using the *constant comparison method* [19]. In this phase each construct pair was compared with each other to decide if they were similar. Further defining and refining of categories occurred as the process was undertaken to produce a final set of thematic tags.

Phase 3 As a final phase, two annotators (one of whom was not involved in Phase 2) used the themes generated in Phase 2 to independently annotate the data. Both of these annotators then discussed areas of disagreement and revised annotations independently. This process is important to reduce idiosyncrasies and improve the reliability of category identification [19, 31]. Agreement between raters was high [Cohen's kappa = 0.76, $p < .001$]. A final discussion took place to resolve remaining disagreements before final annotations were performed by the lead author.

5 Results

Participants across the study produced a total of 266 implicit-emergent construct pairs. The analysis resulted in eight key dimensions¹ that categorise participants' perceptions when comparing dialogue with a human to dialogue with speech-based IPA partners. A summary of the themes, selected construct pairs and total number of constructs linked to each theme are shown in Table 2. Some construct pair examples in the table were reversed so that constructs most closely related to the human partner appear on the left for ease of

¹A miscellaneous category was created when no agreement on category and construct explanation could be traced back to interview data

interpretation (e.g., Human/IPA). A full list of the constructs elicited in each category is included in supplementary material. Below we use data from interview transcripts to explore the definition, and differences between human and speech-based IPA dialogue experiences respectively.

Partner Knowledge Set

72 of the constructs participants generated are focused on the perceived types of knowledge and information that each of the partners can supply during dialogue. Participants said that, unlike humans, speech-based IPAs' knowledge tends not to extend to subjective opinions or judgements. The IPAs' knowledge base was also perceived as more factual, whereas people were seen as relying on more experience based knowledge:

"...they provide very fact and evidenced based knowledge where people will give experience based knowledge" [P22]

On occasions when a speech-based IPA did try to convey a subjective preference, responses were varied:

"Alexa's opinion on Rock music, it said it liked Queen and I think Bohemian Rhapsody...felt very human-like" [P15]

Others felt these interjections made it feel like the IPAs were trying too hard to be human-like, and this served to highlight the role of the team who programmed the IPA to create an illusion of subjective knowledge on a particular theme:

"When Alexa made an X-Files reference I immediately thought of whatever Silicon Valley guy on his lunch break on a Tuesday was like 'Oh, I'm gonna make an X-Files reference'...it was almost trying too hard, it took me out of it"[P1]

Essentially, participants saw IPAs' knowledge as objective "purely informative types of things" [P3] that "just pull up a Wikipedia page...as quickly as possible" [P1]. Conversely, humans were seen as having an ability to draw on richer forms of information meaning they could "talk for like ten minutes" if asked "what do you think of David Bowie?" or on a topic like "not even whether he was a good person but what is a good person"[P1]. This leads to them being perceived as able to give "complex" [P11,P22], and "personalised" [P20,P21] answers to potentially complicated queries. Humans are seen as more able to contextualize, make use of socially relevant and colloquial knowledge, "make assumptions" [P19] and draw supplementary information from the conversation as it unfolds:

"A big part of it is actually talking to someone who knows the greater context...to know the greater context of everything that's going on around you, I guess that's something the virtual assistants just can't do right now" [P10]

Humans were regarded as "spontaneous" [P9], "nuanced" [P14], and "interpretive" [P22]. Yet, although speech-based IPAs were seen as having access to a greater "depth of information" [P5], their "very bookish" nature meant they were perceived as more "constrained" [P5] and "literal" [P9,P22]:

"When I asked what [the human partner] thought about Chuck Norris she was saying like, 'Oh, he's very conservative'. The political aspect came into it...but say if Chuck Norris likes Pizza, well they [the IPAs] wouldn't be able to say then maybe he likes cheese as well, maybe he likes tomatoes as well. They just know he likes Pizza" [P12].

Interpersonal Connection

48 constructs generated emphasized perceived qualities associated with the development of interpersonal relationships between dialogue partners. A number of participants noted that, compared to the interaction with a human, interaction with both Siri and Alexa felt "fake" [P10] or "completely false" [P14]. This was sometimes amplified by "attempts to be funny (and) relatable" [P14], that were interpreted as dry, boring and unfunny and generally seen as an attempt to simulate an artificial social connection. The issue of authenticity of the interaction also contributed to a sense of strangeness or eeriness. One participant remarked that they wanted these systems to have more humanlike qualities, but the eeriness of the current interpersonal interaction was also noted:

"One of the biggest things to me is like 'is this thing humanlike?' And then on the other side I'd say like uncanny...there were times when I felt I was sitting deep in the uncanny valley when I'm talking to these things" [P10]

Some noted a lack of inherent emotion on the part of the IPAs when giving information, whereas it felt that humans generally express inherent emotion about a subject by default when conversing. Again the authentic nature of the interactional qualities were brought into question, where people felt that, although IPAs projected emotion at points, that "they were only pretending" [P22].

Empathy and engagement were also prominently mentioned. IPAs were described as largely devoid of the ability to show warmth compared to a human partner. The IPAs felt "quite disengaged compared to the human" [P24], especially when responding to "personal questions" like "where are you from? And how are you today?" [P24]. Alexa's tendency to "churn out answers" [P24], meant it felt less "relatable" and less "human-like". Likewise the IPAs' "lack of feeling and lack of humor" [P5] made them seem like "disinterested" [P16] dialogue partners. Although some participants did characterise all the partners they interacted with as friendly, there were clear deficits in how relatable and interpersonal the IPAs were compared to the human partner.

Table 2: Summary of construct analysis

Theme	Construct pair examples (Ordered Human/IPA)	Number of construct pairs (% of all constructs generated)	% of agreement between raters
Partner Knowledge Set	Opinionated/Non-judgmental; Biased/Neutral; Free-Bookish; Expansive/Limited; Ad-hoc/Spontaneous/Pre-programmed; Colloquial/Universal knowledge; Abstract/Specific knowledge; Lateral/Inflexible thinking	N=72 (27.1%)	95.8%
Interpersonal Connection	Personal relatability/Manufactured; Genuineness/Ungenuine; Real/Fake; Canny/Uncanny; Emotional/Cold; Personal/Robotic; Connection/Disconnected-disinterested; Engaged/Remote; Humour/Humourless	N=48 (18.1%)	83.3%
Linguistic Content	Short answers/Long answers; Safe/Edgy; Expansive/To-the-point; Convenience/Inconvenience; Elaborate/Pointed; Polite/Blunt or rude; Slang/Phrasing; Colloquial/Formal; Vague/Detailed	N=41 (15.4%)	51.2%
Partner Performance and Capabilities	Channels of communication Multiple/Single; Speed of response: Slow/Fast; Hesitating/Confident-decisive; Recognition-Understandability (High/Low)	N=21 (10.2%)	63.0%
Conversational Interactivity	Two-way/One-way; Conversive/Monologue; Leading/Uninteractive; Conjunctive/Disconnected; Continuation/Stop & Start conversation; Conversational speech (fluid)/Stilted speech; Keeping track/Isolated responses	N=24 (9%)	87.5%
Partner Identity & Role	Humanness/Machineness; Real/Organic-Artificial; Personalised/Commercialised; No agenda/Agenda; Transparency (purpose) (Low\High); To help/To serve; Level standing/Slave	N=21 (7.9%)	100%
Vocal Qualities	Pronunciation/Incomprehensible; Consistent/Choppy; Pleasant/Unpleasant; Audibility-Clarity of speech (High/Low); Enthusiasm/Monotonous; Speech modulation/No modulation; Cheery & Emotional/Dull & Emotionless	N=17 (6.4%)	64.7%
Behavioural Affordances	Natural/Conscious; Familiarity/Unfamiliarity; Patience/Impatience; Easy communication/Limited recognition (make allowances for)	N=13 (4.9%)	84.6%.
Miscellaneous	Fair/Unreasonable; Calculated/General; Personal/Impersonal	N=3 (1.1%)	100%

Linguistic content

41 constructs generated focused on the structure and tone of linguistic content generated by each partner. IPAs' linguistic output was generally seen as formal, "always using full sentences" [P4], along with being "polite" [P20], "diplomatic" [P24] and "safe" [P11]. Between the IPAs, Alexa's use of humor did make it seem less formal than Siri for some participants. However, human-like social content such as small talk was often seen as excessive, inappropriate, "obscure" and "a bit weird" [P19]. Indeed there was a strong mistrust when Alexa's output involved opinion rather than formal fact based content. For instance, one participant stated they felt "manipulated" [P14] when Alexa expressed a favorite band and song.

Participants also felt that human partners used a more narrative rather than descriptive style to convey responses. Although at points, IPAs were perceived as direct and to-the-point in their responses, Alexa was commonly criticized for providing too much information to queries compared to human partners and Siri.

Partner Performance & Capabilities

27 of the constructs were categorised as describing functional aspects of the way dialogue partners performed. Many of these emphasised the multimodal nature of communication and the functional limitations of IPAs. The most significant observations made by participants related to the importance of feedback through non-verbal communication in generating a sense of humanness. Specifically, they felt that communication through speech-only led to ambiguity as to whether their partner had recognised and understood what they had said. Siri's use of visual feedback (e.g. presenting automatic speech recognition (ASR) output on screen) was particularly useful in reducing this ambiguity, and was compared to feedback given from human partners through cues such as facial expressions or hand gestures. Some felt Siri outperformed the human and Alexa in giving confirmation of understanding:

"Well yeah, everything was oral for Alexa, umm, and human, whereas Siri provided visual context as well...it's like a multifaceted approach" [P13]

Others drew similarities between the efficiency of multimodal interactions with Siri and the use of facial and hand gestures:

"She (Human) could give me a physical representation as an answer...something else that I thought was interesting was because you have like the phone, like when it gives you information it doesn't tell you, it just pulls it up...So like people can use various methods but also Siri uses pictures and stuff, whereas Alexa can't do that." [P22]

Despite some issues with ambiguity, recognition and understanding were seen as strong across elements. Yet there were clear differences in how the precision and accuracy of information being delivered by each partner was perceived:

"If I were to take that (Human's) advice I wouldn't have known how to get to there because I wouldn't have known which bus to get, things like that" [P14]

"...the computers because they're automated don't make errors so there's a difference probably in the quality of information" [P24]

The speed of response was also a clear point of divergence between the elements. Many participants noted that the IPAs were very quick to respond, whereas the human was sometimes hesitant:

"...there's like a lead time when you speak to a human...like the, 'Umm' 'Ehh'; where they're thinking like its a pause for thought, whereas a computer doesn't have the pause for thought, so I suppose, speed of response quick" [P24]

Conversational Interactivity

24 constructs related to how conversational participants felt interactions with each of the elements were. A number of the constructs related to perceptions of the direction and interactivity of communication in these contexts with IPAs containing, "very little conversation and more, like commands" [P10]:

"The human yep, very good conversation, very interactive, the others...it definitely didn't feel like it was a conversation. Like I was asking questions and waiting for a response" [P9]

Participants perceived that the questions used in the study, when used with the human partner, were the start of a conversation that could unfold dynamically, rather than merely a request that was completed after an answer:

"With the first question 'how are you today?'; the human, kind of responded with 'yeah I'm good, how are you?'; whereas, both Siri and Alexa were just an answer, and end of conversation" [P3]

"When I asked do you like rock music and she (Human) said, that depends what you mean by rock music, like that's the beginning of a conversation, it's not just an answer" [P1]

This perceived lack of progression with IPA dialogue was consistently mentioned, with participants noting that to get more from the IPA you would have to "...ask them another question" [P20]:

"With the human...when they give an answer, you can go on again. But with them (IPAs) it's like, once they give an answer, that's it" [P18]

Likewise the fluidity of the conversational interaction with IPAs was an issue:

“Some of the things they (IPAs) say is like stilted. Like, for example when she was telling that joke she’s kinda like pausing, but that’s not natural” [P18]

Partner Identity or Role

21 constructs focused on the partner’s identity and the role they play within the dialogue. On the whole these constructs highlighted a simple dichotomy of partner identity between human and machine. There were also reflections on the commercial nature of IPAs and how this influences functionality:

“...I’m aware of the company, because I know these are products sold for money” [P1]

“Like Google Assistant, I probably linked it to Spotify but I’m pretty sure the first time I asked it to play music it was like you can open this on Google Music” [P6]

The power dynamic between user and speech-based IPAs also formed part of the theme’s constructs:

“Me and human are at the same level...when I’m (IPAs) asking the questions I don’t really feel I’m asking the questions, I’m dictating for an answer...it’s a master-slave type thing” [P14]

On the other hand the IPA’s purpose was seen as very “transparent” [P6] compared to humans who may have more complex motives and social roles within interaction.

Vocal Qualities

17 constructs emphasised aspects of non-lexical voice quality. This largely revolved around clarity of a partner’s voice. Participants felt that “*their [IPAs] diction or their actual pronunciation was very sharp*” [P4], and that all partner’s voices were “*...easy to understand. They’re all comprehensible*” [P15]. This highlights improvements in speech synthesis and the development of clear and understandable voices.

Clear differences between partners were seen when discussing voice expressiveness. Humans we were seen to “*express emotions when we talk*” but “*when they (IPAs) say something, they say it really coldly. Whereas...from a human it would have some ups and downs, or like some voice modulations that Alexa and Siri doesn’t have*” [P19].

Some did note the attempt to include expressiveness in the voices used by IPAs, feeling that they were “*...all relatively cheery. They didn’t sound, in any way, dull...(but) the human was much more emotionally cheery I suppose*” [P16].

Behavioural Affordances

13 constructs focused on affordances people felt they needed to make to account for limitations of their dialogue partners. The vast majority of these types of observations related to participants feeling they had to adapt the way they spoke when interacting with speech-based IPAs by trying to

“.. pronounce my words more accurately” due to consciously considering “*whether they understood*” [P21]. Participants also felt that they could be “*more informal, or maybe more spontaneous*” [P9] when interacting with a human, and that queries made to speech-based IPAs were less natural and had to be structured in a particular way:

“I know how to ask a question to get the information that I want that wouldn’t necessarily be the way you would converse with a human” [P10]

This was a particular issue for some who felt interactions with IPAs had rules that needed to be learned:

“I wasn’t familiar with like, how, to interact with them, I didn’t know the rules” [P3]

Finally, some participants suggested they would be more forgiving and patient in accommodating communicative limitations of humans compared to IPAs.

6 Discussion

Research [15, 16, 20, 38, 40, 44] has consistently identified that humanness is fundamental to speech interface design, supporting the metaphor of human dialogue for interaction. It is currently unclear what dimensions are important in users’ conceptualisation of humanness and how people see speech interface functionality through this lens. Through using the Repertory Grid Technique, we discover eight key dimensions which are the focus of humanness: 1) a partner’s level and type of knowledge or *partner knowledge set*; 2) the potential for *interpersonal connections*; 3) the *linguistic content* they use in their dialogue turns; 4) the presence or absence of *conversational interactivity*; 5) functional attributes that relate to a *partner’s performance and capabilities* to complement dialogue; 6) a coarse grained judgement of their *identity and role* in the dialogue, 7) the expressiveness and clarity of their *vocal qualities* in the speech they produce and 8) *behavioural affordances* made to accommodate for perceived limitations of dialogue partners. These dimensions, based on categorisation of user-generated constructs, echo previous researcher led attempts to map fundamental differences between humans and machines [27] (e.g. ‘interpersonal connection’ is similar to ‘emotional responsiveness vs inertness’, and ‘partner knowledge set’ mirrors ‘cognitive openness vs rigidity’).

Although people may use humanness as an anchor [15, 38] people make clear distinctions between humans and speech-based IPAs on these dimensions. Participants see IPAs as being more fact based than opinionated in the knowledge they have, with human partners being able to combine knowledge and ideas in novel ways to contribute to dialogue. Attempts made by speech-based IPAs to be socially relatable and to create a connection were regularly

seen as fake, with IPAs seen as lacking in emotion or perceived interest in the dialogue. IPAs were also perceived as far more formal and verbose in the volume of content they deliver, with social or opinion based content being viewed with suspicion. Differences in the levels of feedback between humans and IPAs were also expressed, whereby speech-only devices were limited when compared to partners that use visual or other multimodal signals to convey understanding. Speech-based IPAs were judged as more able to return accurate and faster responses compared to human partners. Whilst human dialogue was two way and more expansive in nature, IPAs by comparison seemed limited to question-answer type interactions. Above the clear human/machine dichotomy participants also perceived that IPAs were built to serve the user, with humans having more complex social roles and purposes when conversing. Although voice clarity did not vary between partners, there were clearly perceived differences in expressiveness and paralinguistic features, whereby the speech-based IPAs seemed cold and less emotive than a human. Participants also noted a perceived lack of flexibility and a need to understand the rules of interaction when communicating with speech-based IPAs, with humans seen as more adaptable to how the user may want to converse.

The multi-dimensional nature of humanness

Our major contribution is in identifying the key dimensions that users use to conceptualise humanness and how these are mapped to speech-based IPA interactions. This adds much needed richness to the view of humanness as an influencer of the speech-based IPA user experience [17, 38], by breaking down the concepts, identifying important dimensions, and identifying the similarities and differences in how these are applied to humans and speech-based IPAs. From our findings it is clear that humanness needs to be considered not as unidimensional, but as a multi-dimensional concept in speech interface design. The primary purpose of the analysis was not to generate design recommendations, rather it sets to identify the dimensions that influence perceptions of humanness and how they differ in human and IPA interaction. That said, designers could use these dimensions to inform decisions and areas of focus. For instance, our themes could aid in understanding how users frame their experience when compared to human interaction, thereby identifying pros and cons of using humanness as a design metaphor. Striving for humanness is a common guiding principle in speech interface design [3] and a heuristic users frame their experiences through [17, 36, 38]. A nuanced understanding of how users comprehend their interactions through this heuristic is valuable to inform design decisions. With regards to research, a number of

themes do indeed echo topics within speech interface research. For instance, vocal quality clearly maps to work on speech synthesis, where developing more human-like, expressive [1], emotive [39] and personality-filled [51] voices is currently underway. User research has also focused on exploring the role of humanness in partner knowledge assumptions [15], vocal quality [16, 37], partner identity [4], linguistic content [13] and conversational interactivity [13, 45]. Our work not only identifies the important dimensions of humanness when considering user perceptions, but formalises intuitions by researchers on the types of design decisions that may affect these perceptions.

Issues with emulating aspects of humanness

Our findings show that humans and speech-based IPAs are clearly perceived differently on the dimensions identified. IPA interaction is seen as predominantly unidirectional and as a master-servant relationship compared to human-based dialogue interaction. They are also perceived as more socially fake and limited in their interpersonal ability than human partners, as well as seeming more formal and specific than human interlocutors.

This supports the notion that, although they may be perceived through the lens of human dialogue, speech interface dialogues are fundamentally distinct [13, 44]. Currently, dialogue interactions with speech-based IPAs tend to be isolated question-answer pairs [44], seemingly functional rather than interpersonal or social in nature [13]. It may be argued that this is because of truly natural aspects of dialogue not being implemented appropriately or at all in the types of devices currently available. Especially in terms of interpersonal connections and conversational interaction, current projects are exploring how to include this in speech interface design [18, 23, 43, 48] meaning this may become more common and implemented more effectively. Our data suggests some fundamental barriers to accepting such an interaction. Small talk was sometimes flagged as inappropriate or undesirable for current speech-based IPA partners. More social features were perceived as inauthentic, owing to the perceived pre-programmed nature of these systems, rather than because of the types of interactions speech-based IPAs were designed for. Although context and the purpose of the interaction may play a role in the acceptability of social functionality, this type of functionality may be pushing the fundamental limits of spoken interactions between humans and machines [10, 40].

That is not to say that certain dimensions of humanness identified in our work should not be considered in future speech interface design. Concentrating on other dimensions of human-human communication may be more appropriate and even beneficial to user experience for speech interfaces interactions, including speech-based IPAs. In particular

for IPAs, our participant highlighted making minor adjustments towards more human-like ways of delivering linguistic content, like favouring brevity in message delivery and including multimodal feedback available in dialogue. This may improve the interaction experienced. Aspects such as developing expressive vocal qualities may also be useful in more accurately communicating emotive content through synthesis.

7 Limitations & Future Work

Because of the aim of the work, and the triading paradigm inherent to RGT, participants were asked to make direct comparisons between a human and two speech-based IPAs using the human as comparator. This meant that the human was highly salient in the comparisons being made between the three elements, potentially influencing which differences are prioritised. The purpose of this salience was to elicit dimensions and constructs particularly focusing on how humanness may be framing dimensions that people consider in speech-based IPA dialogue. This was so as to reflect findings in the literature that emphasise the important of humanness in design and the role this may have in mental model development [15, 38]. Future work could include more elements to limit the impact that each may play in guiding user constructs.

Our study involved participants reflecting on their interaction with the three elements. This interaction revolved around nine pre-determined questions. This technique was useful to expose participants to particular interaction issues and to ensure all participants could reflect on similar experiences when eliciting constructs. Yet it may have limited the types of dialogues participants reflected on, especially when interacting with the human dialogue partner. Future work may look to use more free-form interactions with all partners, which may not only replicate the categories for constructs found in this research, but add to their richness.

It is also important to note that our study gives a snapshot of user perceptions on these dimensions, yet does not provide information as to how these may be fluid over an interaction, change with different experiences, or how these may change from initial perceptions before interaction. Mapping this dynamism has been noted as a consistent challenge in partner model and perspective taking research, in both human [4] and machine [16] dialogue. Future approaches to this topic should focus on novel methods to gain a view on the dynamics of user perceptions. The themes identified may also act as a starting point to develop a measure to allow researchers and designers to quantify the impact of design choices on the dimensions of humanness. Validated metrics to quantify user perceptions are needed in the field [11]. We plan to use this work as a starting point for developing such a metric.

8 Conclusion

Concepts of humanness are core to the design of speech interfaces like IPAs, yet the specific dimensions of humanness that people use to define these interactions are not fully understood. Our study has clearly outlined key themes related to how users view humanness in dialogue interaction and how this varies in speech-based IPA dialogue. It highlights that whilst humanness may be integral to speech interface design, significant thought needs to be placed into how humanness may be achieved and implemented with sensitivity to the specific dimensions identified.

Acknowledgments

This research was supported by an employment based PhD scholarship funded by the Irish Research Council and Voysis Ltd (R17830).

References

- [1] Kei Akuzawa, Yusuke Iwasawa, and Yutaka Matsuo. 2018. Expressive Speech Synthesis via Modeling Expressions with Variational Autoencoder. In *Interspeech 2018*. ISCA, 3067–3071. <https://doi.org/10.21437/Interspeech.2018-1113>
- [2] René Amalberti, Noëlle Carbonell, and Pierre Falzon. 1993. User representations of computer systems in human-computer speech interaction. *International Journal of Man-Machine Studies* 38, 4 (April 1993), 547–566. <https://doi.org/10.1006/imms.1993.1026>
- [3] Matthew P Aylett, Benjamin R Cowan, and Leigh Clark. 2019. Siri, Echo and Performance: You have to Suffer Darling. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, alt08.
- [4] 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 (Oct. 2011), 41–57. <https://doi.org/10.1016/j.cognition.2011.05.011>
- [5] Cynthia Breazeal. 2003. Emotion and sociable humanoid robots. *International journal of human-computer studies* 59, 1-2 (2003), 119–155.
- [6] Susan E Brennan and Herbert H Clark. 1996. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22, 6 (1996), 1482.
- [7] Allison Bruce, Illah Nourbakhsh, and Reid 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. IEEE, 4138–4142.
- [8] Justine Cassell. 2001. Embodied conversational agents: representation and intelligence in user interfaces. *AI magazine* 22, 4 (2001), 67.
- [9] Vincent Cho and Robert Wright. 2010. Exploring the evaluation framework of strategic information systems using repertory grid technique: a cognitive perspective from chief information officers. *Behaviour & Information Technology* 29, 5 (2010), 447–457.
- [10] 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>, Vol. 76.
- [11] Leigh Clark, Phillip Doyle, Diego Garaialde, Emer Gilmartin, Stephan Schögl, Jens Edlund, Matthew Aylett, João Cabral, Cosmin Munteanu, and Benjamin Cowan. 2018. The State of Speech in HCI: Trends, Themes and Challenges. *arXiv preprint arXiv:1810.06828* (2018).

- [12] Leigh Clark, Abdulmalik Ofemile, Svenja Adolphs, and Tom Rodden. 2016. A multimodal approach to assessing user experiences with agent helpers. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 6, 4 (2016), 29.
- [13] Leigh Clark, Nadia Pantidi, Orla Cooney, Philip Doyle, Diego Garialde, Justin Edwards, Brendan Spillane, Christine Murad, Cosmin Munteanu, Vincent Wade, and Benjamin R. Cowan. 2019. What Makes a Good Conversation? Challenges in Designing Truly Conversational Agents. *arXiv:1901.06525 [cs]* (Jan. 2019). <https://doi.org/10.1145/3290605.3300705> arXiv: 1901.06525.
- [14] Benjamin R Cowan and Holly P Branigan. 2015. Does voice anthropomorphism affect lexical alignment in speech-based human-computer dialogue?. In *Sixteenth Annual Conference of the International Speech Communication Association*.
- [15] Benjamin R Cowan and Holly P Branigan. 2017. They Know as Much as We Do: Knowledge Estimation and Partner Modelling of Artificial Partners.
- [16] Benjamin R. Cowan, Holly P. Branigan, Mateo Obregón, Enas Bugis, and Russell Beale. 2015. Voice anthropomorphism, interlocutor modelling and alignment effects on syntactic choices in human-computer dialogue. *International Journal of Human-Computer Studies* 83 (Nov. 2015), 27–42. <https://doi.org/10.1016/j.ijhcs.2015.05.008>
- [17] 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.
- [18] Laurence Devillers, Sophie Rosset, Guillaume Dubuisson Duplessis, Lucile Bechade, Yucel Yemez, Bekir B Turker, Metin Sezgin, Engin Erzin, Kevin El Haddad, Stephane Dupont, et al. 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)*. IEEE, 697–701.
- [19] Penny Dick and Devi Jankowicz. 2001. A social constructionist account of police culture and its influence on the representation and progression of female officers: A repertory grid analysis in a UK police force. *Policing: An International Journal of Police Strategies & Management* 24, 2 (June 2001), 181–199. <https://doi.org/10.1108/13639510110390936>
- [20] Mateusz Dubiel, Martin Halvey, and Leif Azzopardi. 2018. A Survey Investigating Usage of Virtual Personal Assistants. *arXiv:1807.04606 [cs]* (July 2018). <http://arxiv.org/abs/1807.04606> arXiv: 1807.04606.
- [21] Jens Edlund, Julia Bell Hirschberg, and Mattias Heldner. 2009. Pause and gap length in face-to-face interaction. *Columbia University* (2009). <https://doi.org/10.7916/d82f7wt9>
- [22] Daniel Fallman and John Waterworth. 2010. Capturing user experiences of mobile information technology with the repertory grid technique. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments* (2010).
- [23] 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).
- [24] Fay Fransella, Richard Bell, and D Bannister. 2004. *A manual for repertory grid technique*. Wiley, Hoboken, NJ. OCLC: 53978922.
- [25] Moojan Ghafurian, Neil Budnarain, and Jesse Hoey. 2019. Role of Emotions in Perception of Humanness of Virtual Agents. In *Proceedings of the 18th International Conference on Autonomous Agents and Multi-Agent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 1979–1981.
- [26] Emer Gilmartin, Marine Coltery, 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 Press, Glasgow, UK, 31–32. <https://doi.org/10.1145/3139491.3139494>
- [27] Nick Haslam. 2006. Dehumanization: An integrative review. *Personality and social psychology review* 10, 3 (2006), 252–264.
- [28] Nick Haslam, Paul Bain, Lauren Douge, Max Lee, and Brock Bastian. 2005. More human than you: Attributing humanness to self and others. *Journal of personality and social psychology* 89, 6 (2005), 937.
- [29] Chin-Chang Ho and Karl F MacDorman. 2017. Measuring the uncanny valley effect. *International Journal of Social Robotics* 9, 1 (2017), 129–139.
- [30] Trevor Hogan and Eva Hornecker. 2013. Blending the repertory grid technique with focus groups to reveal rich design relevant insight. In *Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces - DPPI '13*. ACM Press, Newcastle upon Tyne, United Kingdom, 116. <https://doi.org/10.1145/2513506.2513519>
- [31] Devi Jankowicz. 2004. *The easy guide to repertory grids*. Wiley, Chichester, West Sussex, England ; Hoboken, N.J.
- [32] George Kelly. 1991. *The psychology of personal constructs. Vol. 2, Vol. 2.*, Routledge in association with the Centre for Personal Construct Psychology, London; New York. OCLC: 973052807.
- [33] Alan Kennedy, A Wilkes, L Elder, and Wayne Murray. 1988. Dialogue with machines. *Cognition* 30 (1988), 37–72. [https://doi.org/10.1016/0010-0277\(88\)90003-0](https://doi.org/10.1016/0010-0277(88)90003-0)
- [34] 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*. ACM, 1191–1194.
- [35] Ludovic Le Bigot, Jean-François Rouet, and Eric Jamet. 2007. Effects of Speech- and Text-Based Interaction Modes in Natural Language Human-Computer Dialogue. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 49, 6 (Dec. 2007), 1045–1053. <https://doi.org/10.1518/001872007X249901>
- [36] Lucian Leahu, Marisa Cohn, and Wendy March. 2013. How categories come to matter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. ACM Press, Paris, France, 3331. <https://doi.org/10.1145/2470654.2466455>
- [37] Kwan-Min Lee and Clifford Nass. 2005. Social-Psychological Origins of Feelings of Presence: Creating Social Presence With Machine-Generated Voices. *Media Psychology* 7, 1 (Feb. 2005), 31–45. https://doi.org/10.1207/S1532785XMEP0701_2
- [38] 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.
- [39] Elizabeth J. Meddeb and Patricia Frenz-Belkin. 2010. What? I Didn't Say THAT!: Linguistic strategies when speaking to write. *Journal of Pragmatics* 42, 9 (Sept. 2010), 2415–2429. <https://doi.org/10.1016/j.pragma.2009.12.022>
- [40] Roger K. Moore, Hui Li, and Shih-Hao Liao. 2016. Progress and Prospects for Spoken Language Technology: What Ordinary People Think. 3007–3011. <https://doi.org/10.21437/Interspeech.2016-874>
- [41] Masahiro Mori, Karl F MacDorman, and Norri Kageki. 1970. 2012. *The Uncanny Valley [From the Field]*. *Robotics & Automation Magazine, IEEE* 19, 2 (1970), 2012.
- [42] Sharon Oviatt, Jon Bernard, and Gina-Anne Levow. 1998. Linguistic Adaptations During Spoken and Multimodal Error Resolution. *Language and Speech* 41, 3–4 (July 1998), 419–442. <https://doi.org/10.1177/002383099804100409>

- [43] Ioannis Papaioannou, Amanda Cercas Curry, Jose L Part, Igor Shalymov, 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).
- [44] 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.
- [45] 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 - CHI EA '17*. ACM Press, Denver, Colorado, USA, 2853–2859. <https://doi.org/10.1145/3027063.3053246>
- [46] Valentin Schwind, Katrin Wolf, and Niels Henze. 2018. Avoiding the uncanny valley in virtual character design. *interactions* 25, 5 (2018), 45–49.
- [47] Mildred L.G. Shaw and Laurie F. Thomas. 1978. FOCUS on education—an interactive computer system for the development and analysis of repertory grids. *International Journal of Man-Machine Studies* 10, 2 (March 1978), 139–173. [https://doi.org/10.1016/S0020-7373\(78\)80009-1](https://doi.org/10.1016/S0020-7373(78)80009-1)
- [48] 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.
- [49] James C Thompson, J Gregory Trafton, and Patrick McKnight. 2011. The perception of humanness from the movements of synthetic agents. *Perception* 40, 6 (2011), 695–704.
- [50] Lai Lai Tung, Yun Xu, and Felix B. Tan. 2009. Attributes of Web Site Usability: A Study of Web Users with the Repertory Grid Technique. *International Journal of Electronic Commerce* 13, 4 (July 2009), 97–126. <https://doi.org/10.2753/JEC1086-4415130405>
- [51] Mirjam Wester, Matthew Aylett, Marcus Tomalin, and Rasmus Dall. 2015. Artificial Personality and Disfluency. *Interspeech 2015* (2015), 5.
- [52] Jakub Aleksander Zlotowski, Hidenobu Sumioka, Shuichi Nishio, Dylan F Glas, Christoph Bartneck, and Hiroshi Ishiguro. 2015. Persistence of the uncanny valley: the influence of repeated interactions and a robot's attitude on its perception. *Frontiers in psychology* 6 (2015), 883.

This figure "sample-franklin.png" is available in "png" format from:

<http://arxiv.org/ps/1907.11585v2>