Better Curious Than Smart?: Enhance Inclusiveness Between Mismatched Conversational Partners: An Opinion Paper

Guanyu Huang
University of Sheffield
Sheffield, S1 4DP, United Kingdom
ghuang10@sheffield.ac.uk

Roger K. Moore
University of Sheffield
Sheffield, S1 4DP, United Kingdom
r.k.moore@sheffield.ac.uk

CHI'23,, April 23–28, 2023, Hamburg, Germany
ACM 978-1-4503-6819-3/20/04.
https://doi.org/10.1145/3334480.XXXXXXX

Abstract

UPDATED—March 29, 2023. As a rapidly developing user interface, conversational interaction has been adopted in an increasing number of technology products. In theory, the conversational user interface is seen as a tool that can cut through all complexity in human-agent interaction. Still, in reality, it is not yet so effective, especially for non-mainstream users, such as elderly people. One of the reasons could be that such agents and human users are considered mismatched partners in conversations. In this regard, this article illustrates why speech-enabled artificial agents, as mismatched partners, would create barriers to inclusiveness. Adding to that, by looking at other mismatched conversational partners, it seeks ways to enhance inclusivity for human-agent conversational interactions using a collaborative approach.

Author Keywords

Human-Agent Interaction, mismatch partners, inclusiveness, dialogue design, multimodality

CCS Concepts

•**Human-centered computing** → **Interaction paradigms; Interaction design theory, concepts and paradigms;** •**Social and professional topics** → *User characteristics;*

Introduction

Conversation-based interactions between human users and artificial devices have been prevalent in recent years. People have shown great interest in interacting with devices via a conversational user interface (CUI) [17]. As much as people would like to give devices commands by speaking and even interacting with conversational agents (CAs), the real-life adoption and user experience could be disappointing [20] and far from expectations created by popular media [1, p.186]. CUI has hit a bottleneck. Among these difficulties lies inclusiveness, especially the design of CUI. 45% of people aged 70+ felt that 'technology is not inclusive' in 2022 [2]. Facing an ageing society [23], how to make CUI more inclusive to diverse users is a pressing issue.

Understanding how conversational interaction works may help reveal the route for inclusive CUI. Hence, this opinion paper aims to explore inclusive design principles based on the fact that human users and artificial speech-enabled agents are mismatched conversational partners. The article is organised as follows. Firstly, it highlights why CAs, as mismatched partners, would cause inclusivity problems. Secondly, it aims to seek potential ways to enhance inclusivity by looking at other mismatched partners. Finally, it concludes with proposed design principles to develop a more inclusive CUI.

CUI as a Mismatched Partner

From Sounds, Commands to Conversation

Whilst CUI has its convenience when people's hands and eyes are occupied [6], having verbal communication with devices seems not to be so natural. It could be because people never interacted with artificial agents in a natural language before the advent of CUI. Artificial sounds in the past only served as a prompt in human-agent interaction (HAI), such as the 'ping' sound of a microwave oven to indicate the completion of the scheduled time or the joyful sound of an ice cream van to tell you to come out and buy ice creams if you want. The development of CUI has been through several phases: from message receivers and doers of *Command and Control Systems*, to answer providers in *Interactive Voice Response System*, then to a new role as potential emotional supporters in social domains [19].

How Mismatches Affect Inclusiveness

Developing an effective and inclusive CUI in a more dynamic and diverse environment is much more complex. One major reason is the nature of conversations. As grounded in the social need to cooperate [27, 12], conversations require interlocutors to work together to establish common ground [15, p.112][12]. The process of building common ground relies on a shared field of knowledge and experience, abilities to interrelate messages on a moment-by-moment basis and to adjust one's communication efforts according to the environment, another interlocutor(s) and one's own status [24]. It is a continuous mind-reading process to predict and process information [4, 5, 18].

Thus, conversational partners that have different sets of experience, knowledge and skills can be regarded as 'mismatched partners', such as human users and CAs. This mismatch creates barriers to communication.

Firstly, the prior experience, knowledge and skills of a CA are built via rule-based, generative-based or knowledge-baed models [14]. All these enable CAs to make inferences based on the collected information. If users behave in a way that is outside the previously obtained rules or data, the CA has little room to make flexible adjustments. The burden of cooperation then falls on the user. Sometimes due to the lack of clues, users have to explore the boundaries of CA's capabilities and the possibilities of interaction on their own. It could be a steep learning curve for people who lack

Explanantion for Figure 1: Interactive Lemon

Interaction Efforts & Mismatch level: The higher the mismatch level is, the more effort is required to align understandings.

Scenarios: The slope of effort variation depends on interaction scenarios. The slope is gentler for low-demanding scenarios with a large trial-and-error space (e.g. human-pet, parent-infant interactions). For high-demanding scenarios with less room to accommodate errors (e.g. a lost passenger catches a flight in a hurry), the slope of effort is steeper.

Individual threshold: In practice, how much effort one is willing to pay is not indefinite. It varies individually (dotted line). Factors such as interaction needs, motivation, ability and experience can influence thresholds. Individuals with low thresholds are more likely to give up on interaction efforts due to the pressure caused by the mismatch.

experience with conversational technology. The more effort required, the less likely the interaction will be satisfactory. Also, based on Fogg's Behaviour Model (FBM), actions depend on motivation and ability [10]. Thus, if users are not willing or able to put in this extra communication effort, they are not likely to use the CUI and, therefore, would be excluded from the benefits of this type of product.

Secondly, there is another layer of mismatch, that is, the mismatch between a CA's affordance (e.g., its look and voice) and its capabilities. CAs, with or without embodiment, are favoured to have a human-like design (e.g., [3]). According to the affordance theory, people's perception of an object would affect their expectation of it and their behaviours when using it [11, 22, 21, 16]. Thus, counter-intuitively, the type of human-like design may give users an illusion about the actual conversational capabilities of CA, adding to the difficulty in reducing uncertainty in the interaction. This is especially true for users new to CA and/or influenced by science fiction or media.

The way in which mismatches affect interaction is shown in Figure 1.

Potential Ways Ahead for Inclusive CUI

What kind of users would be excluded from CUI? From the above exploration of the spoken interaction, those likely to be excluded are users who share little overlapped experience, knowledge and skills with CUI, who face steep learning curves of CUI and who may be misled by CUI's human-like affordance. Elderly users could fall into one or many of these groups. To include more users in the effective interaction zone (darker area of the 'interactive lemon' shown in Figure 1), CUI designers could aim to reduce the mismatch level, increase error-tolerant space in scenarios and raise the individual threshold of interaction efforts. Some specific

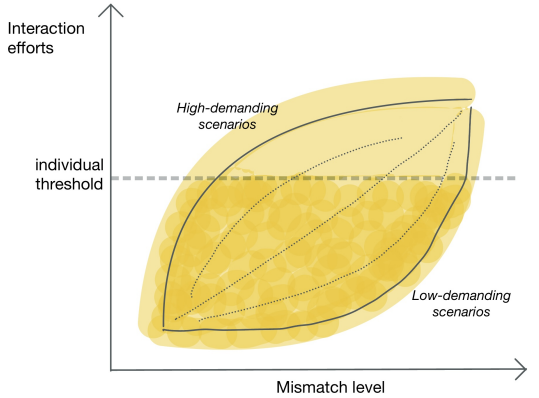

Figure 1: The figure named 'Interactive Lemon' shows how interactive efforts vary with the mismatch level in scenarios.

methods can be found in the following examples.

Lessons Learnt from Other Mismatched Partners
One typical example of mismatched partners' interaction is between first-language (L1) and second-language (L2) speakers. To make spoken interaction work, challenges for L2 speakers are categorised [9] and problem-solving mechanisms are proposed [8]. It is suggested that L2 speakers plan ahead, including their use of language, time to digest information, self-correction and actively seek clarification. L1 speakers can also play a role in improving communication outcomes by various means, such as improving their attitudes towards L2 speakers [26, 25].

HAI can also work with deficiencies in a transparent manner. In addition to improving the skills needed to enhance interaction, it seems sensible to keep users informed about the agent's limits and work with them instead of acting like

a smart human-like agent. In this way, users may feel involved in the development. Affordance design aligned with a conventional agent's abilities could be a good starting point to reduce potentially misled expectations [13]. Also, just like L1 speakers can learn how to talk with L2 speakers, it is worth noting that actionable properties of objects can be constructed cognitively by learning [21]. New products which have similar designs to the old ones can share intrinsic affordances [16]. Thus, having a learning curve is not necessarily a bad thing. The question is how to make learning enjoyable and make these newly acquired interaction methods gradually become intrinsic.

Learn from Other Interactive Interface
Dialogue is a collaborative process. Yet, conversational agents' abilities to adjust their behaviours instantly are limited. Despite this, CA can consider how to involve the other party in this collaborative interaction, especially when facing difficult moments. One technique worth trying is artificial curiosity. For example, a curiosity-driven model has achieved good results [7]. With the right dose of curiosity, it is possible for CA to leave a little more room for clarifications for interlocutors when they need it instead of reacting based on self-interpreted intentions.

Conclusion
To make CUI more inclusive, this opinion paper takes a step back and looks at the theoretical background of why CUI could be a mismatched partner for human users and how such mismatches could affect inclusivity regarding interaction efforts to pay. On the basis of reviewing strategies that have worked for other mismatched partners, it proposes that the potential for relieving the bottleneck and achieving a breakthrough in CUI is in the transparent and collaborative design.

Acknowledgements
This work was supported by the Centre for Doctoral Training in Speech and Language Technologies (SLT) and their Applications, funded by UK Research and Innovation [grant number EP/S023062/1].

REFERENCES
[1] Christoph Bartneck, Tony Belpaeme, Friederike Eyssel, Takayuki Kanda, Merel Keijsers, and Selma Šabanović. 2020. *Human-robot interaction: An introduction*. Cambridge University Press.

[2] Brittne Kakulla. 2021. Older Adults Embrace Tech for Entertainment and Day-to-Day Living. (2021). `https://www.aarp.org/research/topics/technology/info-2022/2022-technology-trends-older-americans.html` [accessed 23-February-2023].

[3] Rodney Brooks. 2002. Humanoid robots. *Commun. ACM* 45, 3 (2002), 33–38.

[4] Andy Clark. 2014. Perceiving as predicting. *Perception and its modalities* (2014), 23–43.

[5] Andy Clark. 2015. *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.

[6] Philip R Cohen and Sharon L Oviatt. 1995. The role of voice input for human-machine communication. *proceedings of the National Academy of Sciences* 92, 22 (1995), 9921–9927.

[7] Malcolm Doering, Phoebe Liu, Dylan F Glas, Takayuki Kanda, Dana Kulić, and Hiroshi Ishiguro. 2019. Curiosity did not kill the robot: A curiosity-based learning system for a shopkeeper robot. *ACM Transactions on Human-Robot Interaction (THRI)* 8, 3 (2019), 1–24.

[8] Zoltán Dörnyei and Judit Kormos. 1998. Problem-solving mechanisms in L2 communication: A psycholinguistic perspective. *Studies in second language acquisition* 20, 3 (1998), 349–385.

[9] Zoltán Dörnyei and Mary Lee Scott. 1997. Communication strategies in a second language: Definitions and taxonomies. *Language learning* 47, 1 (1997), 173–210.

[10] Brian J Fogg. 2009. A behavior model for persuasive design. In *Proceedings of the 4th international Conference on Persuasive Technology*. 1–7.

[11] James J Gibson. 1977. The theory of affordances. *Hilldale, USA* 1, 2 (1977), 67–82.

[12] Thomas M Holtgraves. 2013. *Language as social action: Social psychology and language use*. Psychology Press.

[13] Guanyu Huang and Roger K Moore. 2022. Is honesty the best policy for mismatched partners? Aligning multi-modal affordances of a social robot: an opinion paper. *Frontiers in Virtual Reality* (2022).

[14] Shafquat Hussain, Omid Ameri Sianaki, and Nedal Ababneh. 2019. A survey on conversational agents/chatbots classification and design techniques. In *Web, Artificial Intelligence and Network Applications: Proceedings of the Workshops of the 33rd International Conference on Advanced Information Networking and Applications (WAINA-2019) 33*. Springer, 946–956.

[15] Robert M Krauss and Susan R Fussell. 1990. Mutual knowledge and communicative effectiveness. *Intellectual teamwork: Social and technological foundations of cooperative work* (1990), 111–146.

[16] Sorin Adam Matei. 2020. What is affordance theory and how can it be used in communication research? *arXiv preprint arXiv:2003.02307* (2020).

[17] M McTear, Z Callejas, and D Griol. 2016. The conversational interface: Talking to smart devices: Springer international publishing. *Doi: https://doi.org/10.1007/978-3-319-32967-3* (2016).

[18] Roger K Moore. 2016. Introducing a pictographic language for envisioning a rich variety of enactive systems with different degrees of complexity. *International Journal of Advanced Robotic Systems* 13, 2 (2016), 74.

[19] Roger K Moore. 2017. Is spoken language all-or-nothing? Implications for future speech-based human-machine interaction. In *Dialogues with Social Robots*. Springer, 281–291.

[20] Roger K Moore, Hui Li, and Shih-Hao Liao. 2016. Progress and Prospects for Spoken Language Technology: What Ordinary People Think.. In *INTERSPEECH*. San Francisco, CA, 3007–3011.

[21] Donald A Norman. 2008. Affordances and Design. jnd. org. (2008). `https://jnd.org/affordances_and_design/` [accessed 05-March-2022].

[22] Donald A Norman and Emotional Design-Why We Love. 2004. *Everyday Things'*. Basic books.

[23] World Health Organization. 2022. Ageing and Health. (October 2022). `https://www.who.int/news-room/fact-sheets/detail/ageing-and-health` [Online; posted 01-October-2022].

[24] Teresa Pierce and Amy M Corey. 2009. *The evolution of human communication: From theory to practice*. EtrePress.

[25] Marina Sarda Gou. 2020. *Attitudes towards social robots and how they are affected by direct and indirect contact with a humanoid social robot*. Ph.D. Dissertation. University of Sheffield.

[26] Nicholas Close Subtirelu and Stephanie Lindemann. 2016. Teaching first language speakers to communicate across linguistic difference: Addressing attitudes, comprehension, and strategies. *Applied Linguistics* 37, 6 (2016), 765–783.

[27] Michael Tomasello. 2009. *Why we cooperate*. MIT press.