Leaving the Butler Behind: The Future of Role Reproduction in CUI

Donald McMillan
donald.mcmillan@dsv.su.se
Department of Computer and System Sciences
Stockholm University
Stockholm, Sweden

Razan Jaber
razan@dsv.su.se
Department of Computer and System Sciences
Stockholm University
Stockholm, Sweden

ABSTRACT

Speech technologies are increasing in popularity by offering new interaction modalities for their users. Despite the prevalence of these devices, and the rapid improvement of the underlying technology, how we actually interact with these devices has remained wrapped up in the metaphors of command and control based around the problematic reproduction of the role of butler, maid, or personal assistant. In this paper we explore the issues around focusing our development and research on making a 'better' subordinate, and point to some opportunities to replace and refresh the status quo.

CCS CONCEPTS

• **Human-centered computing** → **Natural language interfaces**; **Interaction design theory, concepts and paradigms**; *Interaction design process and methods.*

KEYWORDS

Speech Interfaces, CUI, Interaction Design

ACM Reference Format:
Donald McMillan and Razan Jaber. 2021. Leaving the Butler Behind: The Future of Role Reproduction in CUI. In *3rd Conference on Conversational User Interfaces (CUI '21), July 27–29, 2021, Bilbao (online), Spain.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3469595.3469606

1 INTRODUCTION

The number of references to Hal from Stanley Kubrick's 1968 film *2001: A Space Odyssey* or the M3-B9 G.U.N.T.E.R robot from the 1960's CBS series *Lost in Space* sprinkled through the publications and presentations on CUIs show that there is a highly compelling popular-culture vision of how systems can be spoken to, and their agency. Yet those designing interaction should note that the behaviours of such fictional conversational user interfaces are defined by them being a protagonist in an ongoing fictional plot. In taking inspiration from such pop-culture tropes, we must also carefully examine what not to take, and adding technical protagonists to our already complex lives – even in the limited sense of how CUIs interact now – must be reasoned and deliberate if it is done at all.

CUI '21, July 27–29, 2021, Bilbao (online), Spain

ACM ISBN 978-1-4503-8998-3/21/07…$15.00
https://doi.org/10.1145/3469595.3469606

We propose that the "retro-tech" [44] default of relying on examples of digital servants, maids and butlers, to situate interaction puts us firmly in the camp of those who, as Robertson put it, "advanced technology in the service of traditionalism" – specifically perpetuating ethnocentrism, paternalism, and sexism.

In this provocation paper we first give a brief overview how such interaction is viewed in the literature, discuss some issues with the current metaphors used in agent based interaction, and suggest different social roles that could be explored in the development of understandable and communicative metaphors of interaction

2 BACKGROUND

While there is a large body of work related to conversational user interfaces. From processing and preparing the incoming audio [19, 22, 52] transcribing what was the user said [17, 62], to understanding what the user meant [19] all have the goal of allowing users to speak naturally and fluently to a system in multiple complex contexts of use. From this point there is development focused on deciding what action the system should take as a result [1, 23, 24, 41, 60, 61], what exactly the spoken response should be [10, 26, 27, 31, 45, 48, 57], and how that response should sound [10, 14, 32, 46, 56, 66].

Taking these together there is some work that looks at the overall interaction with such devices, be that in public areas [20, 40] or in social [6] or home settings [4]. What is even less evident is going beyond using the metaphor of an subservient conversationalist as a way to piggyback on human social conventions to help users understand how they should interact with such a system.

3 ROLE BASED CUIS

There are many ways to interact with computer systems using your voice. In this paper we are not concerned with interaction paradigms which revolve around verbalising commands without the expectation of (much of) a vocal response, or those which monitor how, when, and what people say for language training [11, 29, 35], speech therapy [3, 12, 13, 18, 36, 39], or medial assessment [38, 62] purposes. What we are focusing on here is interaction where an agent is used as an intermediary between the user and the system they wish to interact with, and the social role that agent is mimicking. We do not question the potential scientific or social value in research on agents able to converse, emote, or elicit emotion. Rather, we want to examine the use of these roles in designing an interaction metaphor in a system where the goal of the interaction is to turn on lights or find restaurants. We want to ask: What is the benefit of basing the intermediary agent's interaction on this particular role for the user and the system?

We are also not making the argument that agents should be eschewed in preference for direct manipulation or non-skeuomorphic CUIs. Not only are there many examples of interaction where the strengths of the agent architecture can be used in combination with the strengths of the systems with which it interacts, but as visual interfaces show the move away from skeuomorphism can be slow and in many cases contains echoes of the skeuomorphic metaphors that were once relied on.

Agent interactions give the possibility to employ various techniques observed and defined in Conversation Analysis, using the agent to some extent to take the place of a human interlocutor in ongoing interactions – interrupting or waiting its turn, addressing particular speakers, or keeping up with conversational context. Yet, for these agent interactions to work, an agent must be provided with some sort of behavioural dialogue, actions, and reactions. In this, we see an area where change is necessary – even by those who may see these role-based agents as mere stepping-stones towards non-skeuomorphic CUIs, as these behaviours could potentially echo through CUIs as long as the floppy disc icon echos through GUIs.

3.1 Leaving the Butler Behind

An understandable metaphor helps users frame, remember, and anticipate interactions and can define a field. In HCI the metaphor of the desktop, with files, folders, and a wastebasket, has endured for half a century since being introduced by Alen Kay in Xerox, ahead of the MOAD [15]. While there has been a continual push in HCI to move beyond the desktop [25, 37], the legacy of the file folder, inbox, and outbox on the top of a desk persist. The currently most pervasive metaphor in agent-based CUI is of ordering around a subordinate, a deracialized servant figure [42] in the form of a maid, butler, or *smart wife* [55]. They are an instantly and continuously available when called for and invisible when not, and the majority of the tasks they are able to – or designed to – undertake are traditionally viewed as "housework" taking on roles of domestic workers—who have historically been human women of colour, who were the "invisible absorbers" of the "physical and affective 'dirt' of a home" [47] However, as a research field, we must reflect upon whether this is the metaphor that we want to employ going forward.

The metaphor has, at its root, the problematic human-human relationship of the domestic servant and the 'master' – one which persists across the globe and perpetuates differentiation by race, gender, and class [7, 49]. Yet even as a global metaphor, there are stark differences in the understanding of the role of, and relationship with, a domestic servant between cultures [54, 65]. In comparison to the knowledge workers' desk, relatively well standardised through globalisation of standards and available resources [64], this metaphor may not even hold up to the imperfect understanding of the desktop by users [43].

In question is also the impact of taking such a fraught and complex human-human relationship and necessarily distilling it to the point where it can be simulated. While the anthropological work on the relationships between servants and families shows a great deal of complexity – especially around the relationship to those needing care, such as children [53] and the elderly [8] – the introduction of anthropomorphic virtual agents which can be insulted without recourse and to which polite requests are often less successful than rude ones [6] has caused concern for parents [58]. The abstract understanding that Alexa or Siri are not 'real' is complex, and the expressed worry that this might be training children to interact with others in the same rude, blunt manner prompted both newspaper articles [9] and updates to systems to, if turned on by the parents, admonish insults and ask for politeness [5].

Where spoken language is used to interact with the system without a servant dialogue, some examples we see take a less anthropomorphic approach to the interaction – as Balentine [2] put it, trying to develop a good machine rather than a bad person. The badly personified agent sits between the user and the system, which does not align with the goal of user interfaces being as simple and as natural as possible [59], or that of 'bringing computing machines effectively into the processes of thinking that must go on in "real-time," time that moves too fast to permit using computers in conventional ways' [30]. This begs the question, if in CUI we don't want another 40 years of bad butlers and pedantic personal assistants, what can we do about it?

3.2 Future Directions for CUI Research

One reflex in the face of this may be to attempt to replace the perceived role of the CUI with another one, or even to remove it entirely.

In opening the space for a larger variety of roles to be performed by CUIs there are a number of steps that we can take. One angle is to design CUIs in such a way that there is less of a need to rely on the expectations that the standard roles provide. Shneiderman proposes that we "appreciating the differences between human-human interaction and human-computer interaction" to better understand the cognitive processes surrounding human "acoustic memory" and processing [51]. This could give interface designers the tools be able to integrate speech with more interactions and guide users to successful outcomes without such roles being leaned upon so heavily.

Exploring the expectations of how spoken conversation should work, outside of today's CUIs, can provide a vast range of roles and metaphors that can be drawn on for inspiration [28]. Drawing from conversation analysis, beyond methods for investigating how people use and interact with of CUIs [21, 33], can allow us to apply the formal knowledge of human conversation within turn-taking systems, sequence organisation and repair strategies. Moving towards a multi-turn speech system would provide more opportunity for a dialogue based interaction that shifts from the command-based approach, opening more opportunities for enacting different roles. This analysis of the rhythms and rules of conversational situations [16], in combination with sociology [34], can also provide opportunities for future work to explore the role of both user and speech agent, and bringing the human-centred approach when designing interaction with speech agents. Interaction can be designed to dynamically adapt to the user, cope with changing user behaviours, and improving the underlying models used to understand the user's intent and guide there action as it does so [50]. In this way, roles can be fluid – as they are in human-human interaction – within bounds that can be clearly set and controlled by both designers and end users.

Or one could attempt to remove the problematic role performances from CUIs altogether. However, as noted by Strengers and Kennedy it is best to start from the position "that 'neutrality' is not possible, and queering is difficult, when the very purpose of that robot is to replicate and replace feminized labors." [55] although this does suppose that the CUI will always be an 'other.' With advances in speech production technology providing ever more realistic and computationally tractable options to mimic the voices of others [63] we have the opportunity to generate CUIs that quickly and accurately mirror the prosody and lexical character of the person they are interacting with – and with a little lowering of the pitch it could even match the voice we hear in our heads. While this may sound slightly disconcerting, having something that sounds – and to some extent acts – like your own internal monologue providing reminders or acknowledging commands could provide us the opportunity to explore from another angle the conversation around gender, class, and ethnicity in CUI design.

REFERENCES

[1] Reza Asadi, Ha Trinh, Harriet J. Fell, and Timothy W. Bickmore. 2017. IntelliPrompter: Speech-Based Dynamic Note Display Interface for Oral Presentations. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction - ICMI 2017*. ACM Press, Glasgow, UK, 172–180. https://doi.org/10.1145/3136755.3136818

[2] Bruce Balentine. 2007. *It's Better to Be a Good Machine Than a Bad Person: Speech Recognition and Other Exotic User Interfaces in the Twilight of the Jetsonian Age*. ICMI Press.

[3] Fabio Ballati, Fulvio Corno, and Luigi De Russis. 2018. Assessing Virtual Assistant Capabilities with Italian Dysarthric Speech. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS '18*. ACM Press, Galway, Ireland, 93–101. https://doi.org/10.1145/3234695.3236354

[4] Erin Beneteau, Olivia K. Richards, Mingrui Zhang, Julie A. Kientz, Jason Yip, and Alexis Hiniker. 2019. Communication Breakdowns Between Families and Alexa. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. ACM Press, Glasgow, Scotland Uk, 1–13. https://doi.org/10.1145/3290605.3300473

[5] Edward C Biag. 2018. Kids Were Being Rude to Alexa, so Amazon Updated It. https://www.usatoday.com/story/tech/columnist/baig/2018/04/25/amazon-echo-dot-kids-alexa-thanks-them-saying-please/547911002/.

[6] Michael Bonfert, Maximilian Spliethöver, Roman Arzaroli, Marvin Lange, Martin Hanci, and Robert Porzel. 2018. If You Ask Nicely: A Digital Assistant Rebuking Impolite Voice Commands. In *Proceedings of the 2018 on International Conference on Multimodal Interaction - ICMI '18*. ACM Press, Boulder, CO, USA, 95–102. https://doi.org/10.1145/3242969.3242995

[7] Enobong Hannah Branch and Melissa E. Wooten. 2012. Suited for Service: Racialized Rationalizations for the Ideal Domestic Servant from the Nineteenth to the Early Twentieth Century. *Social Science History* 36, 2 (2012), 169–189.

[8] Elana D. Buch. 2015. Anthropology of Aging and Care. *Annual Review of Anthropology* 44, 1 (Oct. 2015), 277–293. https://doi.org/10.1146/annurev-anthro-102214-014254

[9] CHILDWISE. 2018. *A New Report Shows Children Are Eagerly Embracing Voicerecognition Gadgets – but Could They Be Teaching Children to Be Rude and Demanding?* Press Release.

[10] Nikhil Churamani, Quan Nguyen, Marcus Soll, Sebastian Springenberg, Sascha Griffiths, Stefan Heinrich, Nicolás Navarro-Guerrero, Erik Strahl, Johannes Twiefel, Cornelius Weber, Stefan Wermter, Paul Anton, Marc Brügger, Erik Fließwasser, Thomas Hummel, Julius Mayer, Waleed Mustafa, Hwei Geok Ng, and Thi Linh Chi Nguyen. 2017. The Impact of Personalisation on Human-Robot Interaction in Learning Scenarios. In *Proceedings of the 5th International Conference on Human Agent Interaction - HAI '17*. ACM Press, Bielefeld, Germany, 171–180. https://doi.org/10.1145/3125739.3125756

[11] Gabriel Culbertson, Solace Shen, Malte Jung, and Erik Andersen. 2017. Facilitating Development of Pragmatic Competence through a Voice-Driven Video Learning Interface. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. ACM Press, Denver, Colorado, USA, 1431–1440. https://doi.org/10.1145/3025453.3025805

[12] Harishchandra Dubey, Jon C. Goldberg, Mohammadreza Abtahi, Leslie Mahler, and Kunal Mankodiya. 2015. EchoWear: Smartwatch Technology for Voice and Speech Treatments of Patients with Parkinson's Disease. In *Proceedings of the Conference on Wireless Health - WH '15*. ACM Press, Bethesda, Maryland, 1–8. https://doi.org/10.1145/2811780.2811957

[13] Jared Duval, Zachary Rubin, Elena Márquez Segura, Natalie Friedman, Milla Zlatanov, Louise Yang, and Sri Kurniawan. 2018. SpokeIt: Building a Mobile Speech Therapy Experience. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '18*. ACM Press, Barcelona, Spain, 1–12. https://doi.org/10.1145/3229434.3229484

[14] Alexander J. Fiannaca, Ann Paradiso, Jon Campbell, and Meredith Ringel Morris. 2018. Voicesetting: Voice Authoring UIs for Improved Expressivity in Augmentative Communication. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. ACM Press, Montreal QC, Canada, 1–12. https://doi.org/10.1145/3173574.3173857

[15] H. Fuchs. 1989. An Interactive Display for the 21st Century: Beyond the Desktop Metaphor. In *New Advances in Computer Graphics*, Rae A. Earnshaw and Brian Wyvill (Eds.). Springer Japan, Tokyo, 3–5. https://doi.org/10.1007/978-4-431-68093-2_1

[16] Erving Goffman. 1981. *Forms of Talk*. University of Pennsylvania Press.

[17] Ido Guy. 2016. Searching by Talking: Analysis of Voice Queries on Mobile Web Search. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '16*. ACM Press, Pisa, Italy, 35–44. https://doi.org/10.1145/2911451.2911525

[18] Adam Hair, Penelope Monroe, Beena Ahmed, Kirrie J. Ballard, and Ricardo Gutierrez-Osuna. 2018. Apraxia World: A Speech Therapy Game for Children with Speech Sound Disorders. In *Proceedings of the 17th ACM Conference on Interaction Design and Children - IDC '18*. ACM Press, Trondheim, Norway, 119–131. https://doi.org/10.1145/3202185.3202733

[19] Yuki Hirai and Keiichi Kaneko. 2015. Ambient Conversation Support in Small Face-to-Face Group Meetings. In *Proceedings of the Sixth International Symposium on Information and Communication Technology - SoICT 2015*. ACM Press, Hue City, Viet Nam, 1–8. https://doi.org/10.1145/2833258.2833274

[20] Razan Jaber and Donald McMillan. 2020. Conversational User Interfaces on Mobile Devices: Survey. In *Proceedings of the 2nd Conference on Conversational User Interfaces (CUI '20)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3405755.3406130

[21] Razan Jaber, Donald McMillan, Jordi Solsona Belenguer, and Barry Brown. 2019. Patterns of Gaze in Speech Agent Interaction. In *Proceedings of the 1st International Conference on Conversational User Interfaces (CUI '19)*. ACM, New York, NY, USA, 16:1–16:10. https://doi.org/10.1145/3342775.3342791

[22] Pradthana Jarusriboonchai, Thomas Olsson, and Kaisa Väänänen-Vainio-Mattila. 2014. User Experience of Proactive Audio-Based Social Devices: A Wizard-of-Oz Study. In *Proceedings of the 13th International Conference on Mobile and Ubiquitous Multimedia - MUM '14*. ACM Press, Melbourne, Victoria, Australia, 98–106. https://doi.org/10.1145/2677972.2677995

[23] Hairong Jiang, Ting Zhang, Juan P. Wachs, and Bradley S. Duerstock. 2016. Enhanced Control of a Wheelchair-Mounted Robotic Manipulator Using 3-D Vision and Multimodal Interaction. *Computer Vision and Image Understanding* 149 (Aug. 2016), 21–31. https://doi.org/10.1016/j.cviu.2016.03.015

[24] Shaun K. Kane, Meredith Ringel Morris, Ann Paradiso, and Jon Campbell. 2017. "At Times Avuncular and Cantankerous, with the Reflexes of a Mongoose": Understanding Self-Expression through Augmentative and Alternative Communication Devices. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17*. ACM Press, Portland, Oregon, USA, 1166–1179. https://doi.org/10.1145/2998181.2998284

[25] V. Kaptelinin and M. Czerwinski. 2007. Beyond the Desktop Metaphor in Seven Dimensions. In *Beyond the Desktop Metaphor: Designing Integrated Digital Work Environments*. MIT Press, 335–354.

[26] Yunkyung Kim, Sonya S. Kwak, and Myung-suk Kim. 2011. The Impact of Robots Language Form on People's Perception of Robots. In *Human Centered Design*, Masaaki Kurosu (Ed.). Vol. 6776. Springer Berlin Heidelberg, Berlin, Heidelberg, 253–261. https://doi.org/10.1007/978-3-642-21753-1_29

[27] Tsuyoshi Komatsubara, Masahiro Shiomi, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2014. Can a Social Robot Help Children's Understanding of Science in Classrooms?. In *Proceedings of the Second International Conference on Human-Agent Interaction - HAI '14*. ACM Press, Tsukuba, Japan, 83–90. https://doi.org/10.1145/2658861.2658881

[28] Philip Kortum. 2008. *HCI Beyond the GUI: Design for Haptic, Speech, Olfactory, and Other Nontraditional Interfaces*. Elsevier.

[29] Anuj Kumar, Pooja Reddy, Anuj Tewari, Rajat Agrawal, and Matthew Kam. 2012. Improving Literacy in Developing Countries Using Speech Recognition-Supported Games on Mobile Devices. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems - CHI '12*. ACM Press, Austin, Texas, USA, 1149. https://doi.org/10.1145/2207676.2208564

[30] J. C. R. Licklider. 1960. Man-Computer Symbiosis. *IRE Transactions on Human Factors in Electronics* HFE-1, 1 (March 1960), 4–11. https://doi.org/10.1109/THFE2.1960.4503259

[31] Mike Ligthart, Timo Fernhout, Mark A. Neerincx, Kelly L. A. van Bindsbergen, Martha A. Grootenhuis, and Koen V. Hindriks. 2019. A Child and a Robot Getting Acquainted - Interaction Design for Eliciting Self-Disclosure. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '19)*. International Foundation for Autonomous Agents and Multiagent

Systems, Richland, SC, 61–70.
[32] Nichola Lubold, Erin Walker, and Heather Pon-Barry. 2016. Effects of Voice-Adaptation and Social Dialogue on Perceptions of a Robotic Learning Companion. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, Christchurch, New Zealand, 255–262. https://doi.org/10.1109/HRI.2016.7451760
[33] Donald McMillan, Barry Brown, Ikkaku Kawaguchi, Razan Jaber, Jordi Solsona Belenguer, and Hideaki Kuzuoka. 2019. Designing with Gaze: Tama – a Gaze Activated Smart-Speaker. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 176:1–176:26. https://doi.org/10.1145/3359278
[34] Robert J. Moore and Raphael Arar. 2018. Conversational UX Design: An Introduction. In *Studies in Conversational UX Design*, Robert J. Moore, Margaret H. Szymanski, Raphael Arar, and Guang-Jie Ren (Eds.). Springer International Publishing, Cham, 1–16. https://doi.org/10.1007/978-3-319-95579-7_1
[35] Dania Murad, Riwu Wang, Douglas Turnbull, and Ye Wang. 2018. SLIONS: A Karaoke Application to Enhance Foreign Language Learning. In *2018 ACM Multimedia Conference on Multimedia Conference - MM '18*. ACM Press, Seoul, Republic of Korea, 1679–1687. https://doi.org/10.1145/3240508.3240691
[36] Amal Nanavati, M. Bernardine Dias, and Aaron Steinfeld. 2018. Speak Up: A Multi-Year Deployment of Games to Motivate Speech Therapy in India. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. ACM Press, Montreal QC, Canada, 1–12. https://doi.org/10.1145/3173574.3173892
[37] Nicholas Negroponte. 1989. An Iconoclastic View beyond the Desktop Metaphor. *International Journal of Human–Computer Interaction* 1, 1 (Jan. 1989), 109–113. https://doi.org/10.1080/10447318909525959
[38] Shun Ozaki, Takuo Matsunobe, Takashi Yoshino, and Aguri Shigeno. 2011. Design of a Face-to-Face Multilingual Communication System for a Handheld Device in the Medical Field. In *Human-Computer Interaction. Interaction Techniques and Environments*, Julie A. Jacko (Ed.). Vol. 6762. Springer Berlin Heidelberg, Berlin, Heidelberg, 378–386. https://doi.org/10.1007/978-3-642-21605-3_42
[39] Avinash Parnandi, Virendra Karappa, Tian Lan, Mostafa Shahin, Jacqueline McKechnie, Kirrie Ballard, Beena Ahmed, and Ricardo Gutierrez-Osuna. 2015. Development of a Remote Therapy Tool for Childhood Apraxia of Speech. *ACM Transactions on Accessible Computing* 7, 3 (Nov. 2015), 1–23. https://doi.org/10.1145/2776895
[40] Jennifer Pearson, Simon Robinson, Thomas Reitmaier, Matt Jones, Shashank Ahire, Anirudha Joshi, Deepak Sahoo, Nimish Maravi, and Bhakti Bhikne. 2019. StreetWise: Smart Speakers vs Human Help in Public Slum Settings. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. ACM Press, Glasgow, Scotland Uk, 1–13. https://doi.org/10.1145/3290605.3300326
[41] Bastian Pfleging, Stefan Schneegass, and Albrecht Schmidt. 2012. Multimodal Interaction in the Car: Combining Speech and Gestures on the Steering Wheel. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '12)*. Association for Computing Machinery, New York, NY, USA, 155–162. https://doi.org/10.1145/2390256.2390282
[42] Thao Phan. 2019. Amazon Echo and the Aesthetics of Whiteness. *Catalyst: Feminism, Theory, Technoscience* 5, 1 (April 2019), 1–38. https://doi.org/10.28968/cftt.v5i1.29586
[43] Alisha Pradhan, Leah Findlater, and Amanda Lazar. 2019. "Phantom Friend" or "Just a Box with Information": Personification and Ontological Categorization of Smart Speaker-Based Voice Assistants by Older Adults. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 214:1–214:21. https://doi.org/10.1145/3359316
[44] Jennifer Robertson. 2010. Gendering Humanoid Robots: Robo-Sexism in Japan. *Body & Society* 16, 2 (June 2010), 1–36. https://doi.org/10.1177/1357034X10364767
[45] Frank Rudzicz, Rosalie Wang, Momotaz Begum, and Alex Mihailidis. 2015. Speech Interaction with Personal Assistive Robots Supporting Aging at Home for Individuals with Alzheimer's Disease. *ACM Transactions on Accessible Computing* 7, 2 (May 2015), 1–22. https://doi.org/10.1145/2744206
[46] Najmeh Sadoughi, André Pereira, Rishub Jain, Iolanda Leite, and Jill Fain Lehman. 2017. Creating Prosodic Synchrony for a Robot Co-Player in a Speech-Controlled Game for Children. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction - HRI '17*. ACM Press, Vienna, Austria, 91–99. https://doi.org/10.1145/2909824.3020244
[47] Amy Schiller and John McMahon. 2019. Alexa, Alert Me When the Revolution Comes: Gender, Affect, and Labor in the Age of Home-Based Artificial Intelligence. *New Political Science* 41, 2 (April 2019), 173–191. https://doi.org/10.1080/07393148.2019.1595288
[48] Jan Schneider, Dirk Börner, Peter van Rosmalen, and Marcus Specht. 2015. Presentation Trainer, Your Public Speaking Multimodal Coach. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction - ICMI '15*. ACM Press, Seattle, Washington, USA, 539–546. https://doi.org/10.1145/2818346.2830603
[49] Fiona Schweitzer, Russell Belk, Werner Jordan, and Melanie Ortner. 2019. Servant, Friend or Master? The Relationships Users Build with Voice-Controlled Smart Devices. *Journal of Marketing Management* 35, 7-8 (2019), 693–715.
[50] JongHo Shin, Panayiotis G. Georgiou, and Shrikanth Narayanan. 2013. Enabling Effective Design of Multimodal Interfaces for Speech-to-Speech Translation System: An Empirical Study of Longitudinal User Behaviors over Time and User Strategies for Coping with Errors. *Computer Speech & Language* 27, 2 (Feb. 2013), 554–571. https://doi.org/10.1016/j.csl.2012.02.001
[51] Ben Shneiderman. 2000. The Limits of Speech Recognition. *Commun. ACM* 43, 9 (Sept. 2000), 63–65. https://doi.org/10.1145/348941.348990
[52] Venkatesh Sivaraman, Dongwook Yoon, and Piotr Mitros. 2016. Simplified Audio Production in Asynchronous Voice-Based Discussions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 1045–1054. https://doi.org/10.1145/2858036.2858416
[53] Adéla Souralová. 2015. *Mutual Emotional Relations in Caregiving Work at the Turn of the Twenty-First Century: Vietnamese Families and Czech Nannies-Grandmothers*. Brill. 182–201 pages. https://doi.org/10.1163/9789004280144_009
[54] Lakshmi Srinivas. 1995. Master-Servant Relationship in a Cross-Cultural Perspective. *Economic and Political Weekly* 30, 5 (1995), 269–278.
[55] Yolande Strengers and Jenny Kennedy. 2020. *The Smart Wife: Why Siri, Alexa, and Other Smart Home Devices Need a Feminist Reboot*. The MIT Press, Cambridge, Massachusetts.
[56] Éva Székely, Zeeshan Ahmed, Shannon Hennig, João P. Cabral, and Julie Carson-Berndsen. 2014. Predicting Synthetic Voice Style from Facial Expressions. An Application for Augmented Conversations. *Speech Communication* 57 (Feb. 2014), 63–75. https://doi.org/10.1016/j.specom.2013.09.003
[57] M. Iftekhar Tanveer, Emy Lin, and Mohammed (Ehsan) Hoque. 2015. Rhema: A Real-Time In-Situ Intelligent Interface to Help People with Public Speaking. In *Proceedings of the 20th International Conference on Intelligent User Interfaces - IUI '15*. ACM Press, Atlanta, Georgia, USA, 286–295. https://doi.org/10.1145/2678025.2701386
[58] Alice Truong. 2016. Parents Are Worried the Amazon Echo Is Conditioning Their Kids to Be Rude. https://qz.com/701521/parents-are-worried-the-amazon-echo-is-conditioning-their-kids-to-be-rude/.
[59] Sergej Truschin, Michael Schermann, Suparna Goswami, and Helmut Krcmar. 2014. Designing Interfaces for Multiple-Goal Environments: Experimental Insights from in-Vehicle Speech Interfaces. *ACM Transactions on Computer-Human Interaction* 21, 1 (Feb. 2014), 1–24. https://doi.org/10.1145/2544066
[60] Robert Tscharn, Marc Erich Latoschik, Diana Löffler, and Jörn Hurtienne. 2017. "Stop over There": Natural Gesture and Speech Interaction for Non-Critical Spontaneous Intervention in Autonomous Driving. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction - ICMI 2017*. ACM Press, Glasgow, UK, 91–100. https://doi.org/10.1145/3136755.3136787
[61] Ruolin Wang, Chun Yu, Xing-Dong Yang, Weijie He, and Yuanchun Shi. 2019. EarTouch: Facilitating Smartphone Use for Visually Impaired People in Mobile and Public Scenarios. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. ACM Press, Glasgow, Scotland Uk, 1–13. https://doi.org/10.1145/3290605.3300254
[62] Stephen John Matthew C. Wenceslao and Maria Regina Justina E. Estuar. 2019. Using cTAKES to Build a Simple Speech Transcriber Plugin for an EMR. In *Proceedings of the Third International Conference on Medical and Health Informatics 2019 - ICMHI 2019*. ACM Press, Xiamen, China, 78–86. https://doi.org/10.1145/3340037.3340044
[63] Mika Westerlund. 11/2019 2019. The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review* 9 (11/2019 2019), 40–53. https://doi.org/10.22215/timreview/1282
[64] Elizabeth Yale. 2019. Archives and Paperwork. In *A Companion to the History of the Book*. John Wiley & Sons, Ltd, Chapter 9, 129–142. https://doi.org/10.1002/9781119018193.ch9
[65] B S A Yeoh and S Huang. 1999. Spaces at the Margins: Migrant Domestic Workers and the Development of Civil Society in Singapore. *Environment and Planning A: Economy and Space* 31, 7 (July 1999), 1149–1167. https://doi.org/10.1068/a311149
[66] Serdar Yildirim, Shrikanth Narayanan, and Alexandros Potamianos. 2011. Detecting Emotional State of a Child in a Conversational Computer Game. *Computer Speech & Language* 25, 1 (Jan. 2011), 29–44. https://doi.org/10.1016/j.csl.2009.12.004