

A Workshop on Embodied Vocal Tangible Conversational Agents: a Human Computer Interaction Approach

Mira El Kamali^a, Marine Capallera^a, Leonardo Angelini^a, Omar Abou Khaled^a and Elena Mugellini^a

^aHumanTech Institute, HES-SO//University of Applied Sciences Western Switzerland, Fribourg, Switzerland

Abstract

Tangible User Interfaces (TUIs) and Tangible Interaction are terms increasingly gaining currency within HCI. In order to design a tangible interface, we need to design the digital aspect but also most importantly the physical aspect that will eventually become the new challenge for design and HCI. On the other side, vocal conversational agents are becoming very popular since the development of NLP. Moreover, the market is nowadays producing many vocal assistants such as Apple Siri, Microsoft Cortana and Google Assistant. This workshop focuses on combining the vocal interaction with tangible interaction, thus exploring different embodied vocal conversational agents with tangible aspects. We strive to explore the different tangible designs and the open challenges for embodied vocal conversational agents.

Keywords

Tangible interaction, Conversational agent (vocal), Embodied agent, Physical interface, Human Computer Interaction

1. Introduction

Tangible Interaction is a very interdisciplinary area. It spans a spread of perspectives, like HCI and Interaction Design, and specializes on interfaces or systems that are in how physically embodied - be it in physical artefacts or in environments [1]. Tangible interaction denotes systems that rely on embodied interaction, tangible manipulation, physical representation of data, and embeddedness in real space [1]. Tangible User Interfaces (TUI) were also seen as another way to show graphical displays and that will bring some of the importance of the interaction with physical devices [2]. In particular, combining tangible interaction in IoT devices means that users can send or receive data and information through physical devices. The core idea is to quite literally allow users to grasp data with their hands and to unify representation and control. In fact, researchers are currently exploring the possibilities that arise from the physicalization of digital information especially with the recent development of 3D printing. Stusak

Proceedings of ETIS 2020, November 16–20, 2020, Siena, Italy

✉ mira.elkamali@hes-so.ch (M.E. Kamali); marine.capallera@hes-so.ch (M. Capallera); leonardo.angelini@hes-so.ch (L. Angelini); omar.aboukhaled@hes-so.ch (O.A. Khaled); elena.mugellini@hes-so.ch (E. Mugellini)



© 2020 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

et al. explored the impact of receiving physical artifacts (Activity Sculptures) as a reward of running activities. Results have shown in their 3-week study on 14 people that the 3D printed physical sculptures that represented running data were very useful to motivate the participants to run and to stimulate self-reflection [3]. Tangible interaction can either support immediate interaction in the periphery of the user's attention [4, 5], which is very suitable when users are occupied with other tasks, or it can support meaningful and unexpected interactions that stimulate reflection and the understanding of the system. Tangible manipulation can be direct but also can be often performed without visual attention, relying on proprioception and haptic feedback for evaluating the result of the interaction [4].

Recently, Angelini et al. discussed how the dual nature of tangible interaction and the different properties that it supports can be beneficial for increasing the trust and the user experience with the connected objects, and more in general with the Internet of Things [6].

Conversational Agents are becoming the center of attention in the Machine Learning and Human Computer Interaction area with the development of NLP (Natural Language Processing) [7]. Conversational agents are computer programs that can have a natural conversation in human language. They might be presented to the user in different forms. Conversational agents can come in a form of chat text messaging such as weebot [8] and vocal assistants and vocal assistants such as Apple Siri, Google Assistant. In both cases, the users have a limited understanding of the physical form of the agent. This might be limited to the chatbot avatar, and in the second case only to the voice gender and pitch. Having a full-embodied conversational agent may help the user to increase his/her trust on the agent and to build a stronger emotional bonding[9]. At the same time, the conversational agent embodiment is important to set user expectations in terms of conversation abilities. While users would be happy to discuss and interact with human-like agents, the technology is very often not ready to meet these expectations. In this case, when the agent embodiment is too close to humans, but its behavior or form is not close enough to be perceived by the user as such, the user often develops a feeling of rejection, also known as the "Uncanny Valley" phenomenon[10].

The goal of this workshop was to explore vocal conversational agents embodied in a physical device which we call it "tangible conversational agents (TCA)". This latter should also have a sort of tangibility through direct or tangible manipulation. Typical questions that arise were the following. What are the important properties we need for a tangible vocal device? How to sustain trust and empathy with a tangible vocal device? What are the most properties we need to build a tangible vocal device?

2. Objective

We sought to collect participants' contribution in the field of tangible interaction, tangible user interfaces and conversational agents. During the workshop, the participants were engaged in a brainstorming activity. The objective of the activity was twofold:

- Identify the necessary characteristics that a conversational agent must possess in order

Table 1
Panel schedule

Session	Dur.
Introduction session (Organizers, Context and Objectives)	9 min
Presentation of 2 use-cases (daily life at home and semi-autonomous vehicle)	4 min
Creativity session	15 min
Idea sharing and Discussion	12 min
Closing session	5 min

to qualify as tangible.

- Design a tangible embodied and vocal conversational agent interaction according to a pre-defined situation.

Participants were asked to brainstorm in groups about the concept of having a vocal conversational agent embedded in a physical device with tangible capabilities. The workshop consisted of designing the form of the physical device, the tangibility characteristics and properties of the vocal assistant and the kind of interactions users can have with this embodied conversational agent. In order to center discussion around concrete cases, participants were divided into two groups where each group had a different situation to tackle. The first situation was concerning the use of an agent in daily life at home. The second situation was concerning the use of the agent in a semi-autonomous car. At the end of the panel, we compared the ideas proposed by the 2 groups to see if the agents would be presented with common characteristics, design, interactions or completely different ones.

3. Panel Plan

3.1. Schedule

The workshop has followed a panel format (45 minutes). A tentative plan for the schedule was proposed in Table 1. The organizers begun the panel with an introductory session to detail the organization of the panel and explained the context and objectives of this session. Then, the organizers introduced the creativity session describing the instructions and use-case scenarios. After, there was a brainstorming activity. Attendees were split into 2 groups of 10 to 15 people. Each group worked on a given situation and had to propose a concept at the end of the creativity session. This activity was divided into 2 steps. First, attendees had to identify 2 or 3 characteristics of such an agent (embodied conversational agent) for about 5 minutes. Then, they proposed a very simple design of embodiment and/or a small scenario (about 10 minutes). Finally, each group presented their concept and highlighted the conversational agent's commonalities with the other group working on the other use-case.

3.2. Organization

Because of the current situation, the panel was held fully online. A second Google Meet room were setup for the second group in order to exchange, discuss and debate idea between participants for the creativity activities. The first group might stay on the conference Google Meet room. 2 organizers were present in each room to animate the brainstorming session and answer questions about the situation. A Miro board were also setup for the creativity activity. All participants collaborated on the same board. It was divided into 4 parts- 2 per use-case scenarios: one frame for the characteristics identification phase and another frame for the design. Each group collaborated on the 2 frames corresponding to their use-case.

3.3. Expected Outcomes

The expected outcome of the panel's creativity session were twofold. The first one was a summary of the characteristics and properties of TCA within a specific situation. The second one was a practical insights via the prototypes and scenarios proposed. The comparison of the different agents by groups might allow us to find similarities and explore the differences of a TCA design within different type of situations. Thus, this may put forward the main characteristics needed in any conversational vocal agent to be qualified as a TCA

4. Workshop Results

4.1. Daily Life

During this first part, participants started to think about characteristics the agent could have. Each participant in this group started to put words and images on Miro. Some participants have even mapped different characteristics to then a word that our TCA should be. They mentioned different modalities for different tasks such as:

- vocal: low pitched
- not intrusive
- connected to other things in the house
- shape changing
- not too techy
- proactive, use simple interaction
- soft and squeezable to be hugged → human expression → empathy

Concerning the embodied characteristics, the agent was declared as it shouldn't be intrusive, if the user has guests, then he can choose if he wants to carry this agent. It was also suggested to have a shape changing agent according to user's activity in the house. The agent was declared as well that it should be multi-sensorial and connected to other things in the house. The agent

should be also multimodal for older adults with visual or auditory problems. The agent was also seen as it should be squeezable and with human expression to create emotion and empathy.

Concerning the second activity, participants proposed a small design of such an agent to become a companion to older adults in daily life. As a result, we had 2 main ideas for daily life: (1) A multisensorial system and (2) a companionship robot through its form and being empathic. We agreed that a multi-sensorial system connected in the house that can be also embodied in a conversational agent which is also empathic will encourage the older adult to be more active during his daily life.

4.2. Semi-Autonomous Vehicle

During this first part, participants started to think about characteristics the agent could have. They mentioned different modalities for different tasks such as:

- vocal: sound placement, different volume of voice, voice interaction
- haptic: vibration, temperature, add control button on steering wheel
- smell
- shape changing
- Not just vocal: multimodality

Concerning the embodied characteristics, the agent was seen as it shouldn't be distractive, interact at the periphery, not too big, aesthetic, and eye-free contact. The agent was seen as it should also be adaptive: to driver, the passenger and the context (environment) Concerning the information, the agent could convey, participants mentioned the car self-confidence, guiding. The privacy issue was also mentioned. The agent should be easy to turn on/off, non-listening all the time, propose different modes...

Concerning the second activity, participants proposed a small design of such an agent and also small context of use. The main idea is to use different modalities according to the context to warn the driver and attract his/her attention on the current situation. For example, the agent could use light to indicate the car's self-confidence or it could vibrate. (light and vibration combination as pre-alert was also mentioned). A textured dashboard and shape changing dashboard/wheel could also be an idea to attract driver attention. The agent could also just notify the driver that s/he could continue her/his other activities.

4.3. Final Discussions and Reflections

After the parallel session and the presentation of the results of each use case, we gathered all the results to start grouping similar results from each use case (daily life and semi-autonomous vehicle) to produce a design that can work in different contexts. We concluded that multimodality was essential in both context in terms of interaction with the agent. Shape changing could also be an idea to attract driver attention or user's daily life interaction. Finally, we also agreed together that a multisensorial system is beneficial to control the car or the house, in order to have a better interaction with the user.

5. Short Bibliography of the authors

Mira El Kamali is a PhD student in Informatics at the University of Fribourg, Switzerland and works in collaboration with University of Applied Sciences and Arts Western Switzerland in Humantech Institute. Her research interest focuses on Human Computer Interaction, Conversational Agents and especially machine interfaces for older adults.

Marine Capallera is a PhD student at HumanTech Institute. She is working on conditionally automated driving and she is focusing more specifically on multimodal Human-Vehicle Interaction model for supervision.

Leonardo Angelini is a HCI post-doctoral researcher at the HumanTech Institute and Lecturer at HES-SO (teaching HCI and Machine Learning). He has run several workshops in interaction design, including workshops on tangible interaction with IoT at Ubicomp'16 and CHI'18, workshops on full-body and multisensory interaction at TEI'16 and Ubicomp'16, and a workshop on wearable computing at Automotive UI'14.

Omar Abou Khaled is Professor at the University of Applied Sciences and Arts Western Switzerland. His research fields are Human-Computer Interaction and Wearable and Ubiquitous computing.

Elena Mugellini is head of the HumanTech Institute and Professor at HES-SO (teaching HCI and Machine Learning). She has run several workshops in interaction design, including workshops on tangible interaction at TEI'15 and CHI'18, wearable computing at Ubicomp'13 and Ubicomp'14, conversational agents at Ubicomp'18.

Acknowledgments

The authors would like to thank all the persons who contributed to this paper.

References

- [1] E. Hornecker, Tangible interaction, 2020. URL: <https://www.interaction-design.org/literature/book/the-glossary-of-human-computer-interaction/tangible-interaction>.
- [2] B. Ullmer, H. Ishii, Emerging frameworks for tangible user interfaces, *IBM Syst. J.* 39 (2000) 915–931. URL: <https://doi.org/10.1147/sj.393.0915>. doi:10.1147/sj.393.0915.
- [3] S. Stusak, A. Tabard, F. Sauka, R. A. Khot, A. Butz, Activity sculptures: Exploring the impact of physical visualizations on running activity, *IEEE Transactions on Visualization and Computer Graphics* 20 (2014) 2201–2210.
- [4] E. Hornecker, J. Buur, Getting a grip on tangible interaction: A framework on physical space and social interaction, in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '06*, Association for Computing Machinery,

- New York, NY, USA, 2006, p. 437–446. URL: <https://doi.org/10.1145/1124772.1124838>. doi:10.1145/1124772.1124838.
- [5] S. Bakker, E. van den Hoven, B. Eggen, Peripheral interaction: characteristics and considerations, *Personal and Ubiquitous Computing* 19 (2015) 239–254. doi:10.1007/s00779-014-0775-2.
- [6] L. Angelini, E. Mugellini, O. Abou Khaled, N. Couture, Internet of tangible things (iott): Challenges and opportunities for tangible interaction with iot, *Informatics* 5 (2018) 7. doi:10.3390/informatics5010007.
- [7] H. J. M. Cd, *Advances in natural language processing*, 2015. URL: <https://pubmed.ncbi.nlm.nih.gov/26185244/>. doi:10.1126/science.aaa8685.
- [8] K. K. Fitzpatrick, A. Darcy, M. Vierhile, Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): A randomized controlled trial, *JMIR mental health* 4 (2017) e19. URL: <https://europepmc.org/articles/PMC5478797>. doi:10.2196/mental.7785.
- [9] M. El Kamali, L. Angelini, M. Caon, D. Lalanne, O. Abou Khaled, E. Mugellini, An embodied and ubiquitous e-coach for accompanying older adults towards a better lifestyle, in: *International Conference on Human-Computer Interaction*, Springer, 2020, pp. 23–35.
- [10] M. Mori, K. F. MacDorman, N. Kageki, The Uncanny Valley [From the Field], *IEEE Robotics Automation Magazine* 19 (2012) 98–100. doi:10.1109/MRA.2012.2192811, conference Name: IEEE Robotics Automation Magazine.