Using Lo-Fi Prototyping to Envision Conversational Agents in Public Settings

Arnold Jan Quanjer, Antti Jylhä, Jos P. van Leeuwen
The Hague University of Applied Sciences, The Hague, The Netherlands
a.j.r.b.quanjer@hhs.nl
a.t.jylha@hhs.nl
j.p.vleeuwen@hhs.nl

Abstract: Speech interactions are often associated with virtual assistants and smart home devices, designed primarily for private contexts. A less developed domain is speech interfaces in public contexts. In a smart city development project, we explored the potential of distributed conversational speech interfaces in lampposts. Deploying a research-through-design method, we created a lo-fi prototype of the speech interface that test subjects could interact with during experiments in a lab setting. Our first exploratory prototype consisted of a loudspeaker that acted as the interface and preconceived dialogues designed to investigate the boundaries of desirable and acceptable experiences regarding issues such as privacy. Experiencing the interaction with this rudimentary prototype helped people envision potential use cases and reflect on privacy issues: the dialogues revealed subjective limits of what kind of (personal) information people were willing to share with the lamppost. They also elicited thoughts on possible consequences in the social context of citizens.

Keywords: Smart city, artificial intelligence, conversational agent, speech interface.

1.   Introduction

In a smart city development project, the municipality of The Hague intends to realize a smart city infrastructure in the Scheveningen district by equipping lampposts in the area with sensors and network connectivity. Following a user-centric process, we investigated potential use cases of the infrastructure (van Leeuwen et al, 2018). Field research revealed that residents (1) perceive the municipality of The Hague as distant and believe that it is only moderately informed about what is going on in the neighborhood; (2) experience low social cohesion in the neighborhood; (3) find social networks important; and (4) feel a strong commitment to the quality of life in their neighborhood. Based on these findings, we envisioned including a conversational speech interface in the lampposts, with an A.I. backend to analyze all incoming data for relevance and (communicative) value. This would enable the lampposts to serve as a local social network for citizens and as a communication channel with the municipality to facilitate residents getting heard (van Leeuwen et al, 2018).

However, speech interactions are often associated with intelligent personal assistants (IPAs) and smart home devices (Porcheron et al, 2018), designed primarily for private contexts. Using speech interfaces in public contexts is a less developed domain (Clark et al, 2018) and issues such as privacy and social acceptability become more pronounced than in private contexts. According to Bohn et al (2005), in any ambient intelligence system, such as the lamppost infrastructure, it is important to account for personal and social boundaries in addition to utilitarian value; on a related note, Koelle et al (2018) refer to intruding the social spheres of others. There is evidence that especially infrequent users feel socially embarrassed about using their speech-based assistants with other people present (Cowan et al, 2017; Moorthy & Vu, 2014). Furthermore, the adoption of (and interacting with) IPAs is affected by the transparency of data usage (Cowan et al, 2017) and by how trustworthy the IPA provider is considered by the user (Liao et al, 2019).

Another challenge for the design of new forms of conversational agents is that the expectations of people are largely influenced by existing agents such as IPAs on mobile devices (Clark et al, 2019). According to Moore (2017), this kind of mismatch needs to be overcome in dialogue design of conversational agents to avoid usability problems, while Cowan et al (2017) and Clark et al (2019) suggest to avoid making the agents too human-like but to design human-agent conversations as their own genre.

While it can be hypothesized that an infrastructure of talking lampposts could be used, for example, as a means for citizens to interact with the municipality, publicly interacting with this kind of technology would be unfamiliar to most of the prospective users. Our primary research question is:

What are the boundaries of acceptable experiences regarding issues such as privacy and social desirability of the use of a speech-based conversational agent in public spaces?

Previous research has demonstrated that experience prototyping is a useful technique to uncover user experiences with speech interfaces (Rico & Brewster, 2010). Therefore, we used the 'Wizard of Oz' technique

(Dahlbäck, Jönsson & Ahrenberg, 1993; Oulasvirta et al, 2006) to explore how people would experience this type of interaction and to help them envision potential use cases for the technology. In our research effort, we also contribute our experiences regarding the usefulness of lo-fi prototyping as a means to stimulate test subjects to envision potential use cases and reflect on privacy issues.

2. Research method

We used Research through Design (RtD) as a method to conduct our research. RtD is a form of action-driven research that generates knowledge by creating and evaluating designs and prototypes. In this way "*unique insights are gained through design practice to provide a better understanding of complex and future-oriented issues*" (Godin & Zahedi, 2014). We evaluated an interaction concept by building a lo-fi prototype that allowed test subjects (e.g. local citizens) to experience an engaging interaction that helped them envision potential use cases and reflect on possible social aspects and consequences of the experience.

Figure 1: Lo-Fi prototype used in the laboratory experiment.

2.1 Experiment Setup

Our first exploratory prototype consisted of a loudspeaker on a tripod that functioned as a speech interface (see Figure 1). Following the approach of expandable sequences (Moore, 2018) we prepared and wrote out a number of possible dialogues between the lamppost and citizens and used a text-to-speech engine to simulate and trigger the lamppost's responses in dialogues with test subjects. We conducted experiments in a lab setting. The dialogues were designed to investigate the boundaries of desirable and acceptable experiences regarding issues such as privacy, with progressively more intrusive questions from the lamppost (e.g., upload a photo, what is your name, give your phone number, etc.).After the experience with the prototype participants were interviewed, by means of a fixed set of questions, about their expectations regarding the usefulness and desirability of such a talking lamppost in their neighborhood.

Experiments were conducted inside the university building of THUAS and at Wijkoverleg Scheveningen-Dorp, a community center in the Scheveningen district. Both sessions were held in a confined space. Participants were recruited randomly at the respective locations. In total 12 test subjects participated in the experiment, aged between 24 and 69.

3. Results

During the follow-up interviews, participants expressed interest in a broad variety of use cases for a conversational speech interface in their neighborhood. They also mentioned social opportunities and hindrances.

Social:
- Some participants expressed their concern about being perceived as informants by other residents;
- Still others would be interested to see the lamppost act as intermediary to establish new contacts in the neighborhood;

- A talking lamppost functioning as a meeting point can help to find like-minded neighborhood residents in order to undertake activities or collectively improve the quality of life in the neighborhood ("Cleaning up the park together.").

Privacy:
- Some stated that it might feel 'creepy' if the lamppost would initiate the conversation;
- Being addressed directly with your own name raises questions about privacy ("If the lamppost knows me, I would like to know what he knows about me;" "It should not get too personal;" "I would be very suspicious and dislike it.");
- Younger participants appear to be more critical on privacy issues than older participants;

Communication:
- When the lamppost is explicitly presented as a representative of the municipality, participants expressed that it would be perceived more trustworthy;
- If the lamppost is convincingly presented as a representative of the municipality, some participants would be more inclined to answer relevant questions about the neighborhood, but the usefulness of these questions must be clear ("Now that I understand the usefulness, I will cooperate.");
- All participants find it no problem to share photos with the municipality, the conversation may also be shared with the municipality.

Safety:
- Some participants also found that being addressed directly was something positive ("You are known and that can also increase a sense of security or reduce loneliness;" "It is nevertheless a form of attention, even if it's a machine;" "A 'buddy' who knows I'm there when I have to walk through the darkness in the evenings.");
- Others perceive the lamppost as an easily accessible means to report calamities, accidents and crime.

3.1 Using a lo-fi prototype as a research method
Experiencing the interaction helped people envision potential use cases and reflect on privacy issues. Closed scenarios turned out to be well predictable and were usually fully performed with a satisfactory result. More open conversations were more unpredictable, which made it difficult to execute via the 'Wizard of Oz' method. Improvising answers and typing on the spot works, but this disrupts the flow of the conversation and test subjects had to wait longer for an answer.

4. Discussion
Experimenting with the lamppost prototype has given a broad insight not only into the usefulness of a conversational agent in public spaces but especially into the diversity and nuance in public acceptance.
The experiences with the lo-fi prototype elicited insightful conversations with the test subjects and brought up social and communication aspects as well as privacy and safety issues. The reactions were diverse and nuanced, even if the experience with the prototype was rather simple and brief. It appears that the subjects' mental model of the talking lamppost is a key factor in its acceptability and perceived usefulness. This mental model – what is the lamppost's role in the neighborhood? – is largely influenced by the way the lamppost presents itself as a representative of an organization, e.g., the municipality, the local community, or a commercial stakeholder.

According to some interviewees, mainly younger participants, the functionality of a talking lamppost may have little added value compared to their smartphone. The local, static nature of a talking lamppost should therefore be a distinguishing factor. Compared to a smartphone, a static conversational agent in the public space should add value in different ways:
- The lamppost should be recognizable as a touchpoint of the municipality, as this will increases visibility of the municipality in the neighborhood;
- It is a conversation partner that is always available;
- The lamppost stands out from its surrounding and can therefore be easily found;
- A lamppost can report about quality of life in the neighborhood, with specific locally collected information, and can be easily accessed without the hassle of undertaking a complicated search.

The research team experienced that the use of a lo-fi prototype as a research tool was an effective way of initiating meaningful conversations with test subjects, e.g., about privacy and ethics. It also prompted effective

discourses on the acceptability of this new technology. We therefore argue that lo-fi prototyping is cost-effective and informative tool in researching acceptability.

Although our research question on the boundaries of acceptable experiences with speech interfaces in public environments has not yet been fully answered in this work in progress, but the chosen method of using Wizard of Oz experiments has been demonstrated to be efficient and effective, leading to relevant and insightful results and more in-depth conversations with test-subjects than could be expected from less experiential research techniques.

In our follow-up research we intend to move towards analytic experiments to evaluate acceptability of concrete applications of speech interfaces in public spaces. We will install a functional prototype in a semi-public space to investigate this and also to evaluate how the presence of other people and environmental noise affects the user experience, Furthermore we will try out new use cases to explore their desirability and the social acceptance of interacting with a speech interface in public spaces.

References

Bohn, J., Coroamă, V., Langheinrich, M., Mattern, F., and Rohs, M. (2005) "Social, economic, and ethical implications of ambient intelligence and ubiquitous computing", *Ambient intelligence*, pp. 5-29.

Clark, L., Doyle, P., Garaialde, D., Gilmartin, E., Schlögl, S., Edlund, J., Aylett, M., Cabral, J., Munteanu, C., and Cowan, B. (2018) "The State of Speech in HCI: Trends, Themes and Challenges", *arXiv preprint arXiv:1810.06828*.

Clark, L., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., and Wade, V. (2019) "What Makes a Good Conversation?: Challenges in Designing Truly Conversational Agents", In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, Glasgow, UK.

Cowan, B. R., Pantidi, N., Coyle, D., Morrissey, K., Clarke, P., Al-Shehri, S., and Bandeira, N. (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*, Vienna, Austria.

Dahlbäck, N., Jönsson, A., and Ahrenberg, L. (1993) "Wizard of Oz studies—why and how", *Knowledge-based systems*, Vol. 6, No. 4, pp.258-266.

Godin, D. and Zahedi, M. (2014) "Aspects of research through design", In *Proceedings of DRS 2014: Design's Big Debates*, *1*, Umeå, Sweden.

Koelle, M., Boll, S., Olsson, T., Williamson, J., Profita, H., Kane, S., and Mitchell, R. (2018) "(Un) Acceptable!?!: Re-thinking the Social Acceptability of Emerging Technologies", In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (p. W03, Montreal, Canada.

van Leeuwen, J.P., Quanjer, A.J., Jylhä, A., et al. (2018) "Kunstmatige Intelligentie in de Publieke Ruimte in Scheveningen – Projectrapportage", De Haagse Hogeschool, The Hague.

Liao, Y., Vitak, J., Kumar, P., Zimmer, M., and Kritikos, K. (2019) "Understanding the Role of Privacy and Trust in Intelligent Personal Assistant Adoption", In *International Conference on Information in Contemporary Society*, Washington, DC, USA.

Moore, R. K. (2017) "Is spoken language all-or-nothing? Implications for future speech-based human-machine interaction", In *Dialogues with Social Robots*, pp. 281-291.

Moore, R. J. (2018) "A Natural Conversation Framework for Conversational UX Design", In *Studies in Conversational UX Design*, pp. 181-204.

Moorthy, A. E. and Vu, K. P. L. (2014) "Voice activated personal assistant: Acceptability of use in the public space", In *International Conference on Human Interface and the Management of Information*, Heraklion, Crete, Greece.

Porcheron, M., Fischer, J.E., Reeves, S., and Sharples, S. (2018) "Voice Interfaces in Everyday Life", In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI '18), Montreal, Canada.

Oulasvirta, A., Engelbrecht, K.P., Jameson, A., and Möller, S. (2006) "The relationship between user errors and perceived usability of a spoken dialogue system." *ISCA/DEGA Tutorial and Research Workshop on Perceptual Quality of Systems*, Berlin, Germany.

Rico, J. and Brewster, S. (2010) "Gesture and voice prototyping for early evaluations of social acceptability in multimodal interfaces." In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, Beijing, China.