

Mid-term Report

About us

Our group consists of Rajani Shrestha, Mattias Ovesson, and Stian Rustad. All of us are on our first year on our masters in Informatic: Design, use, and interaction, two of us took our bachelor's degree in Informatic: Design, use, and interaction at the University of Oslo whilst one of us has a bachelor in Digital design & innovation from Halmstad University.

Area of interest

We are interested in the way humans interact with a virtual personal assistant (PA) such as Google assistant and also the similarities and differences between these and if they were interacting with a human assistant. Looking at the capabilities of the assistance and when and how people use the assistance, it's also interesting to think about where they use the assistance. Letting users know what the virtual personal assistance is capable of helping them with.

Voice assistants are virtual agents that understand human speech and are able to respond to questions (Hoy, 2018). Today these virtual assistants (VA) exist in most sold smartphones, laptops and computers (Tulshan & Dhage, 2019). These virtual assistants have many different types of functionalities (Tulshan & Dhage, 2019). People use Siri and Google assistant quite often to perform several tasks such as setting an alarm, asking about weather and we get an instant response. But have you ever thought, VA would be capable of doing more than that if you have been probably using VA till today.

Thus we are doing a comparative research study on how one will react with VA in contrast to a human based on empirical research. Our target groups will be students of informatics at University of Oslo, who are well known with the term virtual assistance and might have used it in some ways.

Background

We have seen an increase in the use of voice assistants in everyday life. They are now easily accessible for everyone with a smartphone, through the use of Siri, Alexa and Google assistant. Previous work in this area has tried to explore what motivates and limits the use of VAs in everyday life and what to consider in future design interactions such as Luger & Sellen (2016). Currently the most current types of use are relatively simple, such as checking the weather or setting reminders (Luger & Sellen, 2016). The study done by Luger and Sellen (2016) in 2016 showed that users used natural/colloquial language at first (such as, 'should I take an umbrella today'), but changed into more simple language if the voice

assistant did not understand the input. It would be interesting to do a similar study today, since these voice assistants get better every year thanks to new updates and more information input by users.

The value of using a voice assistant rather than direct manipulation like in a graphical user interface, is to enable multitasking and using the hands for other activities such as cooking or bicycling (Luger & Sellen, 2016). Another value is the time-saving aspect, since completing tasks with a GUI sometimes requires multiple steps. In the process of learning how their voice assistant works, users removed complex words, reduced the number of words used, and used more specific terms to accomplish their task (Luger & Sellen, 2016).

Previous work in communication between human-robot has been done by Hill et al. (2015). Interesting insights from this article is the fact that language skills transfer easily to human-chatbot communication but there are notable differences in the content and quality of those conversations, and there is an increase of profanity used when talking to a chatbot (Hill et al., 2015). Porcheron et al. (2018) argues that the conversational interaction should be considered *embedded* in a conversation and the voice assistant is not a conversationalist.

Questions

Q1: How does the speech pattern of a human change while speaking to a digital PA in comparison to speaking with a human?

Q2:

Methods

It is a comparative study because where we will be studying the similarities and differences between behavior of humans mainly focusing on the speech pattern, speech tone, use of words(formal , informal) while communicating with VA and humans. To remark these differences we need to conduct a detailed study, collect data and analyze the data. Empirical research is one of the approaches which supports all the requirements mentioned and emphasizes evidence for further investigation. As empirical research is based on observation, the main methods we will be using to collect evidence will be obviously observation. The main focus of the observation will be to carefully observe, record and analyze the behaviour of the participants while they are communicating with Google Assistant and humans.

In addition, we will use the article by Hill et al. (2015) as an inspiration. Hill et al. (2015) analyzed changes in communication when people were talking with an intelligent agent as opposed to another human. In this case, they used a chatbot to represent the intelligent agent, however, we will use the google voice assistant. In Hill et al. (2016) article they analyzed the conversation in seven dimensions, such as words per message, words per conversation, and emoticons used.

In our study, we will give the participant different tasks to complete surrounding the theme of "Travel". The task to complete will be the same for the human-human interaction as for the human-voice assistant, this is to analyze the differences. We will record the entire interaction to have the possibility to transcribe the conversation. With a transcript, we will be able to identify and compare the differences. Other interesting differences such as body movement and gestures will be noted down. An example of a task would be: "You will travel to London in February. Ask about relevant information for your trip."

With a transcript we will analyze five dimensions (1) words per question, (2) words per conversation, (3) did they manage to complete their task, (4) total time consumption, and (5) the attitude: polite, normal or rude.

Observation

The observation will be done at the Escape Cafe in the Informatic Building of UIO. We will request the students at the cafe to take part in our one of the tasks. While the student performs the given task it will be recorded and one of us will take notes of important incidents, gestures and observe them.

Data collection: Voice recording and transcribe

Tools: Google virtual assistance (English)

Participants: Minimum 5 for each task

Location: Escape Cafe, Informatic Building ,UIO

What we are going to analyze

1. Words per questions
2. Words per conversation
3. Did they manage to complete their task
4. Total time consumption
5. Way of communication with VA : more politely, normally, rudely

Tasks for participants to complete

1. Communicate about a random topic with one of us
2. Communicate about a random topic with voice assistant (Google Assistant)
3. Ask to set alarm , set timer, find a restaurant

Theme: Travel

Goal: You will travel to London in february. Ask about relevant information for your trip.

Ex:

- Weather
 - Potential answer: The average temperature is 9 degrees
- How expensive is it to travel to London
 - The average daily price is around 145 pound.
- Currency / rate
 - 1 pound is approximately 12 norweigan krone
- Find activities and attractions
 - London Eye, London bridge, Buckingham Palace
- Nearby cities to visit
 - Brighton, Southampton, Cambridge
- Shopping
 - St.james Street ,
 - Westfield London, Covent Garden, Brent Cross shopping centre
- Cuisine
 - Fish and chips
- Flights

Theme: Furnishing /Interior

Goal: Imagine that you just moved and you have a room that is completely empty of furniture. Your goal is to find furniture for your room with the help of google assistant.

Ex. find a chair, table/desk, couch, lamps,

Theme: Latest Technology

Goal: Buy a new phone / Find information about a new gadget

Ex: Latest Iphone, Mobile for certain budget , Best reviewed smart clock

Findings

We performed a pilot test of our tasks and found out that the task is too vague and needs to be worked on more to get the desired information form the task one of the things that seemed to be a problem was the clarity in the role of the human assistant in the human-human interaction. There were some problems with what they felt like they needed to know this might be remedied with changing the location they are traveling to so that they have as little information as possible to begin with.

In our human to human pilot test there were many problems for our participants to understand their task. In our next iteration we will be more clear about the task Informing the participant.

In our human to machine interaction the participant completed the task by asking questions about the weather and booking a cheap flight. We expected the participant to ask more questions, such as activities to do and how expensive it would be.

Observation from communication to VA :

Very precise and short questions ,

VA suggested to click a link instead of talking that interrupted the communication

Observation from communication to Human :

Communication building from one topic ex: suggestion for hotel, then hotel under budget then cost

Questions were descriptive

Future plans:

Brief and understandable explanation of task they need to perform and what are the objectives of our study

A good explanation for human to human interaction for example (ex: more informal, friendly conversation)

Give a task to explore about a unfamiliar country or place

The one who is supposed to be Human

References

- Hill, J., Ford, W. R., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior*, 49, 245-250.
doi:10.1016/j.chb.2015.02.026
- Hoy, M. B. (2018). Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Medical Reference Services Quarterly*, 37(1), 81–88.
<https://doi.org/10.1080/02763869.2018.1404391>
- Kawulich, B. B. (2005, May). Participant observation as a data collection method. In *Forum qualitative sozialforschung/forum: Qualitative social research* (Vol. 6, No. 2).
- Luger, E., & Sellen, A. (2016). "Like Having a Really Bad PA". *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*.
doi:10.1145/2858036.2858288
- Porcheron, M., Fischer, J. E., Reeves, S., & Sharples, S. (2018, April). Voice interfaces in everyday life. In *proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1-12).
- Tulshan, A. S., & Dhage, S. N. (2019). Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa. In S. M. Thampi, O. Marques, S. Krishnan, K.-C. Li, D. Ciunzo, & M. H. Kolekar (Eds.), *Advances in Signal Processing and Intelligent Recognition Systems* (pp. 190–201). Springer. https://doi.org/10.1007/978-981-13-5758-9_17

Tilbakemelding på gruppe 6 sitt "Wonder document":

Star:

- ★ Fin forklaring av area of interest.
- ★ Gode spørsmål som knyttes godt opp mot area of interest.
- ★ Artikkene dere har funnet virker spennende, nyttige og veldig relevante for temaet.

Wish:

- Vi skulle ønske dere kunne utdypet spørsmålene deres mer. Kanskje forklare litt rundt hvorfor dere ønsker å finne svar på akkurat dette.
- Vi tenker også at det kunne vært fint med litt flere ressurser og litteratur. Spesielt med tanke på at dere tenker å bruke literature review som metode.

Generelt fin innlevering og temaet deres virker som et godt utgangspunkt for det som kan bli et spennende prosjekt.