IntraVox: A Personalized Human Voice to Support Users with Complex Needs in Smart Homes

Ana-Maria Salai[1], Glenda Cook[2], and Lars Erik Holmquist[1]

[1] School of Design, Northumbria University, Newcastle upon Tyne, UK
[2] Nursing, Midwifery and Health, Northumbria University, Newcastle upon Tyne, UK
{ana-maria.salai,glenda.cook,lars.holmquist}@northumbria.ac.uk

Abstract. We present IntraVox, a novel voice-based interaction system that supports users with complex needs by introducing a highly personalized, human voice command between smart home devices and sensors. Voice-enabled personal assistants (e.g., Amazon Alexa, Google Home) can be used to control smart home devices and increase the quality of life. However, people suffering from conditions such as dementia or learning disabilities and autism (LDA) can encounter verbal and memory problems in communicating with the assistants. IntraVox supports these users by verbally sending smart home control commands, based on the sensor data collected, to a voice-enabled personal assistant using a human voice. Over 12 months, we conducted 7 workshops and 9 interviews with 79 participants from multiple stakeholder categories - people with LDA, older and frail people, social and health care organisations. Results show that our solution has the potential to support people with complex needs and increase independence. Moreover, IntraVox could have the added value of reinforcing the learning of specific commands and making visible the inner workings of a home automation system, by giving users a clear cause-and-effect explanation of why certain events occur.

Keywords: Home Automation · Smart Homes· Virtual Assistants.

1 Introduction

Smart home technologies can increase the quality of life of many user groups. For instance, voice-enabled personal assistants (e.g., Amazon Alexa, Google Home) are becoming increasingly popular [24]. These devices have great potential to support people of all abilities and across the life course to live independently at home and optimize their quality of life [44]. In addition to setting reminders, playing music, and providing information, voice-enabled personal assistants can control various technologies in a home such as light, temperature, and blinds.

Despite being popular and affordable, certain sectors of the population can encounter issues in interacting with such devices [44]. One issue, particularly relevant for people with speech impairments, memory problems, and cognitive issues, is the verbal requirement to articulate or recall keywords and syntax, such as "*Alexa/Hey Google, turn on the living room lights*".

In collaboration with a housing association, we undertook a research project to understand how to overcome these issues and support the organisation's customers. Over 12 months, we conducted workshops and interviews with various stakeholder categories - end-users, social and healthcare professionals. Based on their feedback, we developed IntraVox. IntraVox ***verbally*** sends smart home control commands to a voice-enabled personal assistant (e.g., Amazon Alexa and Google Home, but it could easily be generalized to other platforms) using a ***human voice*** based on a contextual situation. For instance, when light sensors detect a decrease in the light intensity, IntraVox can enable Alexa to turn on the lights and close the curtains by playing the recording of a human voice (Figure 1). Our system thus has the potential to reveal the inner workings of home automation, giving users a clear cause-and-effect explanation of why certain events occur. This approach also has the potential of improving the usability of voice-enabled personal assistants by removing the necessity for the user to remember and pronounce specific commands whilst also reinforcing the syntax and usefulness of those commands for future situations (e.g., for when the user wants to turn on the lights themselves). The current implementation of the system is composed of a dedicated computer (Raspberry Pi 4B) to which we attached environmental, touch, proximity, and motion sensors. Using a speaker, the system sends prompts and reminders to householders and commands to a voice-enabled personal assistant which in return controls smart home devices.

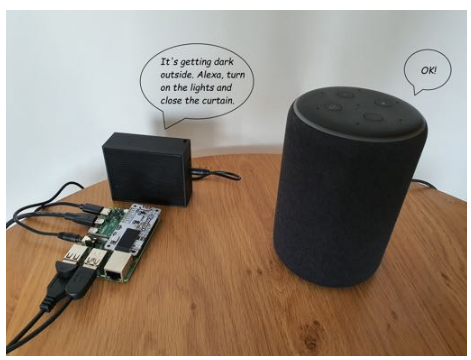

Fig. 1. IntraVox is composed of a Raspberry Pi 4B, sensors, speaker, Amazon Alexa and/or Google Home controlling smart home appliances.

This paper is structured as follows. We start by providing an overview of smart homes and voice-enabled personal assistants, followed by the user sessions conducted to develop and evaluate IntraVox. We then discuss the contributions we bring to the design of Human-Computer Interfaces for smart homes and Ubiquitous Computing.

2 Background

In this section, we outline the current research in the area of smart homes and voice-enabled personal assistants. We present how voice-enabled personal assistants can support independent living and the issues some users suffering from various disabilities encounter when interacting with such devices.

Smart home technologies are designed to support people in living comfortably and independently and have caught the attention of researchers working in the field of Human-Computer Interaction and Ubiquitous Computing [1, 52]. For example, Dahmen et al. [19] developed a digital memory notebook application for a smart home to help older adults keep track of their daily to-do lists. The application, linked with motion and door sensors, keeps track of the activities performed, and provides prompts when remaining activities need to be completed. By using distortions in the wireless signal, Adib et al. [1] created a system that monitors vital signs such as breathing and heart rate. Rafferty et al. [46] retrofitted an assisted living smart home for older people with activity and environmental sensors to create an intention recognition mechanism based on activity data. Daher et al. [18] used in-depth images extracted from a Microsoft Kinect and accelerometer data to determine falls and to localize and track the daily activities of older people. Suitable for studying the behavior of older and disabled people, Li et al. [31] used a less invasive method and developed an indoor human tracking system based on Wi-Fi signal strength. Vines et al. [56] investigated how home technologies can improve the quality of life of older people by developing a non-invasive monitor system composed of sensors connected to a central hub.

Voice-enabled personal assistants have been introduced in the home to facilitate access to information and control smart home instruments. These devices can also support independent living and improve the quality of life of people suffering from various conditions [44]. Pradhan et al. [44] investigated Alexa's accessibility and how it can improve the quality of life of people with motor, sensory, and cognitive impairments. Alexa was predominantly used for listening to music and playing audiobooks, searching for information (e.g., weather forecast), and home automation (e.g., lights, thermostats). Most users with visual impairments and/or motor disabilities mentioned an increase in independence and quality of life because it saved effort and allowed them to communicate with other family members. Alexa also proved useful in helping people with speech impairments to talk slowly, clearly, and loudly. Setting reminders, medication management, managing calendars, and to-do lists was perceived as useful for people with memory difficulties. Finally, the increase in independence also reduced caregivers' responsibilities. Kowalski et al. [28] conducted workshops with older adults to determine the potential of Google Home in the context of smart homes. Results show that older users face barriers (fear of malfunction) but also benefits (intuitive interaction, friendly manner, no handling of devices) when interacting with the assistant. Simpson et. al [50] designed a conversational agent embodied as a household potted flower to provide companionship to older adults living alone. Using an open source framework for the conversational interface, the device engages in a casual conversation with the user regarding a past memory in their life or by suggesting activities that might be of interest to them.

Despite the above mentioned benefits, some user categories encounter challenges in interacting with voice-enabled personal assistants. For example, Atefi et al. [4] discovered that the length of most Alexa Skill commands is greater than 4 words, with some even requesting 8 words. This has led to a complaint by users who tend to find the commands long and difficult. Vacher et al. [55] evaluated a context-aware voice interface system with seniors and people with visual impairments in an Ambient Assisted Living laboratory. The results show that participants disliked the strict grammar rules for the commands and would have preferred a more natural approach to the interaction. Similarly, Sayago et al. [48] argue that the need of knowing the commands and the *one-size-fits-all* approach adapted has an impact on the user experience. Luger and Sellen [33] investigated users' views on Conversational Agents (e.g., Siri, Cortana) and how this affects everyday use. Results show there is a gap between people's expectations and the agent's capabilities. Users started by interacting with the agent in a natural language but were later forced to limit themselves to just keywords and specific terms to obtain the desired output. Moreover, they had to learn how to speak to the agent to be able to successfully interact with it. Similar to previous findings [20], some users were willing to research in advance the agent's capabilities to maximize the interaction and increase trust. Intuitively, investing time in understanding a system's capabilities and learning how to interact with it becomes bothersome, especially for a person with limited cognitive abilities. Porcheron et al. [43] investigated how families engage with Alexa in their homes. The authors developed a voice recorder composed of a Raspberry Pi and a microphone that would record what happened before and during the conversation with the device. Based on the data from a month-long deployment, the authors argue that voice user interface devices should not be viewed as conversational agents since they are not able to properly understand the meaning of a conversation. The authors suggested that voice-enabled personal assistants should be designed based on a request-response rather than a conversation interaction.

The need to speak clearly and loudly with voice-enabled personal assistants is challenging for people with speech impairments [44]. Moreover, the short amount of time in which the commands need to be spoken and the difficulty in remembering these commands constitutes a problem for people with cognitive impairments [33, 44]. Carroll et al. [12] designed and developed an assistive mobile and web platform to support the daily living of people with dementia, by providing voice prompts and task guidance to the users. Being developed as a Skill (a built-in Alexa capability), the system was designed only for people suffering from early stages of dementia as it requires the users to recall specific keywords and phrases. Corbett and Weber [14] developed a mobile voice user interface for people with motor impairments. Regarding speech recognition error rates, the authors highlight that a limited command vocabulary is more effective and accurate than a conversational style interaction. However, from a user experience point of view, the conversational style is the desired and preferred one.

3 Development Process

To address the limitations some users with complex needs face when using voice-enabled personal assistants, we developed a system to support them when using such devices. We developed the system through a process of requirements gathering, system development, feedback interviews, and workshops with social and healthcare organizations and potential end-users. To the best of our knowledge, the resulting *IntraVox* system is unique in that it uses a human voice for Computer-Computer Interaction that aims in increasing the quality of life and independence of people suffering from complex needs (e.g. people with LDA, frail and older people) in the context of a smart home. This interaction focuses on removing the necessity for users to remember specific syntaxes required to enable and control a voice-enabled personal assistant.

3.1 Requirements Gatherings

We started by conducting 2 workshops with 19 Community of Practice (CoP) members from various end-user representation groups such as social care and health organizations, working with and supporting people with LDA. Similar to [32], the workshops aimed to understand how technology can support people with LDA to live independently in their homes. To drive design decisions, we introduced 2 Personas as this User-Centered Design tool proved successful in facilitating design discussions and encouraging participants to engage in a study [10]. The Personas, designed in collaboration with an occupational therapist, represented characters suffering from LDA and behaviours of concern, with reduced independence and significant sensory needs.

During the first workshop (with 11 CoP members), members engaged in timeline analysis to explore the requirements the Personas would have during the day, and how these could be fulfilled through sensing, processing, and actuation. During the second workshop, members engaged in the exploration of challenges and behaviours in relation to daily activities in different locations of a home. The data collected were analysed using elements of thematic analysis [22].

Findings The outcome of these workshops led to 4 concepts where technology could promote users' independence and increase quality of life: (1) **Concept 1: Smart Home** – the concept of automating the house based on sensors (e.g., sound, light, humidity, temperature, and motion) and use of appliances data (e.g., fridge, kettle, microwave, and toothbrush), (2) **Concept 2: Skills Development and Maintenance** – supporting people to develop personal skills (personal care, interests, and activities), (3) **Concept 3: Prompts and Reminders** – provide people with emotional, nutritional (e.g., cook a meal), daily management support (e.g., appointments and daily activities), and (4) **Concept 4: Behaviour and Environment Monitoring** – allow health professionals access to users' activity data. We used these concepts to drive the cycles of development of the IntraVox system.

3.2 System Explorations

The technologies we used as a starting point (e.g., Alexa as a voice-enabled personal assistant, sensors, smart devices) were shaped by the feedback received during the workshops. The challenge was to integrate the different concepts in a system that was easy to use for our potential user groups. To address Concept 1 and 4, we started by designing a system composed of motion, temperature, sound, and light sensors connected to a Raspberry Pi computer, widely used for home automation. Based on the sensor data collected, the Raspberry Pi would send prompts and reminders (Concept 2 and 3) and activate different smart home commands (Concept 1) such as turn on the lights, close the curtains, play music. A voice-enabled personal assistant was selected to carry out the commands as this is a readily available technology and is compatible with a variety of smart home appliances. Similar to [28, 56], it can act as a hub that enables connectivity from different resources and devices. We initially developed 2 ways of sending the instructions in a testing house provided by the housing company partner:

System 1 - *Silent System*: We wanted to have an option where the sensors would control the voice-enabled personal assistant directly without any audible commands. However, Amazon Alexa is fundamentally a voice-enabled personal and home assistant that only accepts verbal commands in its syntax to control devices. This is seen as a shortcoming by some developers, who have created various workarounds. One common way of making it a silent process would require putting a set of headphones on the Amazon Alexa [2]. This is a rather inelegant solution which we did not pursue. Additionally, it can be achieved by using Home Assistant [21], a Python program that can track, control, and automate various smart home devices, and Bespoken [8], a third-party website that creates a virtual/silent Amazon Echo device.

System 2 – *Synthesized Voice System*: Messages verbalized by the Raspberry Pi using voice synthesis created by text-to-speech libraries [47] is another method of enabling Alexa. In our demonstrations, we used an English female voice with emphasis on capitals and speaking slowly. As an example, based on the light sensor data collected, the Raspberry Pi can verbalize a prompt to Alexa to trigger certain processes. Whilst we tried at the beginning to use complex sentences such as "*It is getting dark outside. Alexa, turn on the lights.*" or double commands "*Alexa, turn on the lights and close the curtains.*", we soon realized it is hard for Alexa to understand these commands, regardless of accent and delivery pace of the command. Therefore, we limited the instructions to basic keywords and phrases such as "*Alexa, turn on the lights.*".

3.3 Prototype Interviews

Throughout the development phase, we also held 2 interviews with 3 CoP members who could not attend the workshops (one member being a person with a learning disability). The interviews took place in the same house provided by the housing company to ensure that participants can perceive how the systems would fit in a real home. The aim was to confirm the technology decisions taken based on the feedback collected from the 2 workshops. Participants provided positive feedback with respect to the 2 systems and provided suggestions on how they can be improved. Our CoP member suffering from a learning disability noted that the silent system might not be appropriate for a person with LDA as they could find it "*scary*" and the synthesized voice can be perceived as unpleasant and frustrating. Based on their suggestions we developed IntraVox, a human voice-based interaction system.

4 IntraVox

4.1 Proof-of-Concept

As previously described, the synthesized voice used in System 2 for vocalizing the pre-configured messages can be perceived as unpleasant and difficult for the voice-enabled personal assistant to process. The development process took place in the same house provided by the company. We started by creating audio files with a human voice verbalizing the pre-configured messages and used these to send the prompts, reminders, or instructions to Alexa. The audio files were recorded directly on the Raspberry Pi using a microphone and the Audacity software [5]. The voice could be a voice recording of a relative, a carer, or any other familiar person. The intention was to create a sense of security and comfort, as if the person was there to help, rather than commands being carried out automatically by an anonymous computer.

Python software was written to access each sensor's library, read the sensor values and trigger events accordingly based on the user's preferences. As an example, the user, carer or family member can specify an ideal room temperature (e.g., 22 degrees). Based on *if* statements, the Python software would detect whether the temperature sensor values went above or below the ideal value, and would trigger an event, e.g. play an audio file already saved on its internal memory, using Python libraries for playing audio files. IntraVox is agile and these values can easily be changed by simply modifying the lux (for light), decibel (for sound), degrees (for temperature), etc. values in the Python code. Since we were using a human voice and we were not facing any memory or time length constraints, we decided to expand the messages and adapt them to address different scenarios, as follows:

Home Automation: Mennicken et al. [35] highlight that householders do not like smart homes to make decisions for them. The authors suggest that homes should propose rather than impose certain house processes. To address Concept 1 and to promote independence and empower people to make their own decisions, we adapted the messages to the following syntax (based on light sensor data collected): "*It is getting dark outside. Maybe you should turn on the lights and close the curtains.*". However, this type of message would not be appropriate for people lacking the ability to process this information. We, therefore, designed additional messages such as "*It is getting dark outside. Alexa, turn on the lights and close the curtains.*" (Figure 2).

Based on temperature sensor data, the system also alerts users whether it's getting too warm or too cold in a room and to subsequently, turn on or off the heating or open and close a window. This type of message provides a clear justification for why certain types of home automation processes occur.

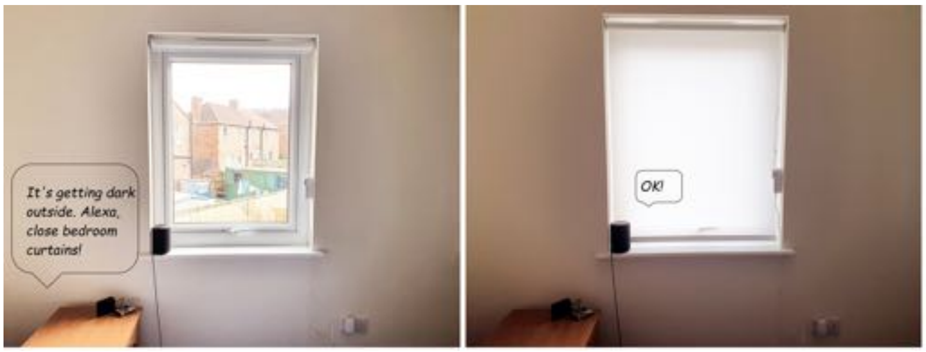

Fig. 2. IntraVox sends instructions to Alexa to close the bedroom curtains.

Activity: Brewer et al. [9] discuss how voice-enabled personal assistants can support the process of remembering by reminding users about activities that need to be performed in the future and by recalling past information. To address Concept 2-4, if no motion has been detected for a certain amount of time, we designed IntraVox to verbalize a reminder such as: "*You should be a bit more active, it's good for you. How about going for a walk?*". On the other hand, if the motion sensors detected increased movement which could be a proxy indicator that the user is agitated, IntraVox would verbalize instructions to Alexa to alleviate the situation (based on the user's preferences): "*You seem agitated. Alexa, turn living room lights purple and play calming music.*".

As identified by Kiseleva et al. [27], successful task completion and low level of effort in seeking information are key factors that contribute to users' overall satisfaction when interacting with a voice-enabled personal assistant. Cho et al. [12] also show that interacting with Alexa requires a pre-decision process and a certain level of energy that users view as a burden especially after a tiring day at work. Therefore, since a user (with or without an impediment) is not required to put effort into remembering and vocalizing a command, our solution could contribute to an enhanced user experience.

4.2 First Validation Workshop

We arranged 2 workshops with 2 stakeholder categories - CoP members and potential end-users - to validate IntraVox and evaluate whether our choice of technologies suits end-users. The workshops took place in the same house provided by the company to ensure that participants can perceive how the systems would fit in a real home.

We first conducted a workshop with 9 CoP members from various councils, social care, and health organizations that work with and support people with LDA, to evaluate the technologies deployed. CoP members took part in interactive in-situ demonstrations of the Silent, Synthesized Voice, and IntraVox systems (Table 1 illustrates the IntraVox concepts both stakeholder categories interacted with during the workshop). They also engaged in technical discussions about sensing and actuation, and provided informed feedback about the systems developed based on open-ended questions.

Table 1. IntraVox concepts both stakeholder categories interacted with during the workshop.

Scenario	Activity
Home Automation	Using light sensors, IntraVox enables Amazon Alexa to turn on/off lights and blinds.
Home Automation	Using temperature sensors, IntraVox prompts users to turn on/off the heating and close/open the window.
Activity	Using motion sensors, IntraVox prompts users to be more active or provides alternatives if they are agitated.

We also organized a workshop with 4 end-users with mild LDA (3 males and 1 female), customers of the housing association that owned the house. The aim was to understand whether the systems developed would be suitable for people with LDA. During the workshop, 4 support workers, employees of the housing association, were also present to offer end-user support. Meetings were held prior to the workshop with one of the support workers to (1) ensure that our workshop structure is appropriate and does not cause any emotional distress in the end-users, and (2) design an easy read version of the information sheet and consent form handed in prior to the workshop.

The workshop lasted for 2 hours and was divided into 2 phase: ***(1) Phase 1 -*** Since no end-user interacted with an Alexa device before, the workshop started by allowing them to familiarize themselves with the device. Similar to [26], end-users were asked to engage with Alexa by conducting short tasks. The tasks varied from asking for cooking recipes, e.g., "*Ask Alexa how to make a sandwich by saying 'Alexa, I want to make a cheese sandwich'*", to asking Alexa to play funny sounds or booking appointments, ***(2) Phase 2 -*** End-users were presented with interactive in-situ demonstrations of the systems (as CoP members) and were asked to provide feedback based on open-ended questions. Notes were taken down and data collected were analysed using elements of thematic analysis [22].

Findings Our preliminary results show that sensors, IntraVox and Alexa are desired by both stakeholder categories (CoP members and end-users and their support workers). Several themes emerged from the discussions held, as follows:

Interaction. Since a voice-enabled personal assistant such as Alexa is part of the IntraVox system, it was important to understand end-users' views on the device and how they interact with it. Confirming previous findings [44], although people suffering from various disabilities can benefit from using an Alexa, their support workers were not aware of the device's capabilities: "*I have Alexa, but I have no idea how to use it, I only play music.*".

Ease of use: Despite not being familiar with the device, the end-users were able to interact with it and understand its responses when undertaking basic tasks such as booking an appointment. More complex interactions (e.g. asking for cooking recipes) were perceived as overwhelming and hard to follow: "*She talks too much.*". Confirming previous research findings [44], processing the command of a person with a voice impairment proved problematic for Alexa. For example, one end-user was addressing the device with "*lexa*" in a quick way and found it frustrating when the system did not respond.

Satisfaction: Purington et al. [45] show that personification is associated with an increased level of satisfaction. People who refer to the device using personal pronouns such as *her* and *Alexa*, rather than *Echo* or *it*, tend to have a more friendly interaction with the virtual personal assistant. Our end-users were using personal pronouns such as *she* to describe their interaction with the device and *Alexa* to enable it. Whilst interacting with the device caused frustration in some end-users due to speech recognition errors, all end-users were satisfied with Alexa. Moreover, our end-users tended to have a polite interaction with Alexa using phrases such as "*Alexa, could you please* [..]?" and "*Thank you.*" once the command was executed. This is similar to the findings by Nass and Brave [37] who show that voice within technology can encourage social interaction and that people respond to a voice socially even if the voice is synthetic.

Voice. The concept of a voice enabling Alexa received positive feedback from both stakeholder categories. For example, end-users appreciated the Activity Scenario and found it to be "*canny*" (nice). CoP members also expressed interest in such technology in the health system: "*We certainly see a place for products like this within social care.*" City council members taking part in the workshop, already trialing the use of Alexa with their customers, also emphasized the system's novelty: "*Interesting to see technology interacting with one another, had initially thought these would have stand-alone operations. Importance of personalising where possible the voice used for the instruction.*".

Silent Communication: Previous Human-Robot Interaction research indicates that humans would like to understand what the robots are communicating with each other using voice and gestures [25]. This in return shows humans that robots can communicate with them, making the interaction feel "*natural and smooth*" as during Human-Human Communication. Williams et al. [57] also show that silent communication between robots is perceived as "*creepy*" by humans and highlight that human-robot settings should be sensitive to human expectations. Similarly, with regards to System 1 – Silent System, end-users' feedback varied from "*I wouldn't be bothered*" to "*I'd get frightened as if it was haunted. What's going on? Like a horror movie.*". Additionally, end-users mentioned they would prefer knowing why certain home automation processes occur: "*Good idea to hear what will happen, to warn me.*".

Personification: All end-users preferred IntraVox to the Synthesized Voice System: "*I want a voice that I recognize.*", highlighting the benefits of having such a system: "*A voice that is familiar calms me down.*". This corresponds to the design paradigm that "*Computers are Social Actors*" and that human-computer relationships are in essence social [39]. Indeed, "*Computers that talk*" aim to create "*natural*" and "*real*" experiences, and the use of a human voice can contribute to an enhanced user experience [32] and increased usability [10]. Previous research also indicates that social robots talking in a human voice is preferred to a synthesized voice and a robotic, monotonous voice might not be suitable for emotional interactions [7].

Trust: The results of the workshop also show that designers might also wish to consider the aspect of trust [20] and that importance should be given to the person the familiar voice belongs to and their relationship with the end-users. Large and Burnett [29] investigated how drivers' attitudes towards GPS (automated driver support system) voices and the impact the voices have on trust and attention. Results show that driver's attitudes are influenced by the "*personality*" of the navigation system. For example, drivers may be less likely to follow the instructions provided by an annoying voice but might follow the ones provided by an entertaining voice. Indeed, our end-users provided mixed feedback regarding what type of voice they would prefer: "*I would like Vicky's* [caregiver] *voice.*" or "*My voice also.*" but "*Not me mam's* [my mother's] *voice*!". It is worth mentioning we used a female voice for the recordings, belonging to one of the workshop's facilitators. Participants also had time to familiarize themselves with the voice before the demonstrations.

4.3 Second Validation Workshop

The findings from the workshops indicated that IntraVox has great potential in supporting people with complex needs. However, the results also show that Alexa might not be suitable for some user categories, as it does not always understand commands, leading to frustration and irritation. This allowed improving the system in cases where users find it challenging to use.

We conducted an additional workshop with 10 CoP members in the same house to understand how IntraVox's functionalities can be expanded and whether it can also improve the quality of life of older adults. As in previous workshops, participants were presented with 2 Personas that were carefully designed by a collaborator in this work, expert in the needs of older people, and an occupational therapist. The Personas represented characters living with long-term health problems (e.g., frailty, dementia, incontinence). CoP members were asked how IntraVox can support the Personas in the context of a smart home, by reflecting also on the 4 Concepts previously discovered.

Findings CoP participants emphasized the importance of personalization and developing solutions based on the user's needs and preferences. They stressed that the needs of older people should be considered alongside the needs of support workers. At all stages of later life, the needs and preferences of older people change and therefore, it is important for the technology to adapt and respond to different requirements. Following their feedback and in collaboration with the expert in older people's needs, various test scenarios were developed to address common challenges encountered by older people - insufficient hydration and urinary incontinence, prompting and maintaining personal hygiene, etc.

4.4 Integrated System

Given the feedback received, we expanded IntraVox by including additional functionalities and sensors (touch and proximity). We also tested IntraVox on a Google Home device to demonstrate the system's compatibility with other voice-enabled personal assistants available on the market. We, therefore, redesigned the messages to include the specific syntaxes and keywords required to interact with the device, e.g. "*Hey Google, close the curtains.*". To address the test scenarios arising from the workshop findings, the following additional IntraVox prototypes have been implemented (in the same test house):

Activity: Socially assistive robots can support people with dementia by prompting interventions when a behaviour change is detected [16] and motivate them in exercising using personalized messages [17]. We, therefore, expanded the previous Activity Scenario to engage users with home-based exercises. Based on the user's daily routine and necessities, whenever inactivity is being detected, the Raspberry Pi would send an instruction to Alexa to start a workout routine: "*You haven't been active in a while. You should work out, it's good for you. Alexa, start 7 minutes workout* [Alexa Skill]." (Concept 3 and 4).

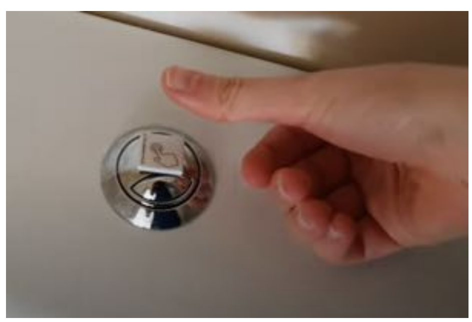

Fig. 3. IntraVox prompts users to wash their hands after using the toilet.

Hygiene and Dehydration: Using audio and picture prompting proved successful in enhancing the independence of adults with Autism Spectrum Conditions by improving their daily living skills such as washing dishes and doing laundry [49]. To addresses Concept 2 and 3, using a touch sensor attached to a toilet (Figure 3), IntraVox can send a verbal notification to the users to wash their hands after using the toilet: "*Remember to wash your hands after going* *to the toilet*.". Dehydration is a common problem in later life, and it can lead to serious health consequences. Since users must drink sufficient fluid, sensors were deployed to determine the amount of water a person is drinking in a day. If the fluid intake is less than 1.5 litres per day [13], a verbal reminder was sent to users to drink the recommended daily intake of fluids. We attached a touch sensor to a glass and an additional temperature sensor at the bottom of a mug that indicates a drink/hot drink is being consumed. In addition, IntraVox can send a message in the morning, afternoon, and evening as a reminder to drink sufficient fluid throughout the day.

Reminder: Using a proximity sensor attached to a front door, users can receive a verbal reminder to take their keys before leaving the house (Concept 2 and 3). Using cloud communications platforms such as Twilio [54] and Plivo [42], IntraVox was programmed to also send an SMS alert to a carer or family member whenever an unexpected or relevant event would occur (e.g., front door opened).

4.5 Online Evaluation Workshops

Following the worldwide lockdown due to the COVID-19 pandemic, evaluating the new IntraVox functionalities during face-to-face workshops was no longer possible. Similar to other research [6], we continued to engage with the same stakeholder categories (potential end-users and CoP members) through internet multimedia platforms and conducted online workshops and interviews.

We organized two online workshops with 15 CoP members and 5 small group interviews involving 14 individuals supporting older and frail people and people with mental health issues. We presented participants with videos of the systems developed (filmed in the same test house) and asked for feedback using open-ended questions. Notes were taken down and data were analyzed using elements of thematic analysis [22]. The workshop focused on understanding how IntraVox can help increase the quality of life of these user categories. In order to validate previous findings, participants were also presented with videos of the Silent and Synthesized Voice Systems.

We also interviewed a frail older man together with his live-in daughter and son-in-law acting as carers, and an older woman together with her daughter who also acted as a carer. Similar to the end-users with LDA workshop, the interviews lasted for one hour were divided into 2 phases: ***(1) Phase 1 -*** End-users visualized videos that compared the interaction between Alexa and Google Home, e.g., asking both devices for cooking recipes. Previous studies evaluated customer satisfaction of Alexa and Google Home [8, 12]. However, to the best of our knowledge, studies have not been conducted to determine which device is received as better by older people, and ***(2) Phase 2 -*** End-users were presented with videos of the concepts developed (as CoP members) and were asked to provide feedback based on open-ended questions. Notes were taken down and data collected were analysed using elements of thematic analysis [22].

Findings Interest was shown by individuals who had not previously considered voice-enabled personal assistants or who have used the assistants in a limited way. There was a consensus among CoP members that products such as sensors and voice-enabled personal assistants have great potential to augment care and support individuals with complex needs to live independently: "*Lots of applications for the technology to support how people may live, manage a long term illness or recover from a health episode.*". Several themes emerged from the discussions held, as follows:

Interaction. Similar to the end-users with LDA workshop, it was important to understand end-users' views on both voice-enabled personal assistants since they are part of the IntraVox system.

Ease of use: Similar to the end-users with LDA workshop findings, the frail older man and his family members recommended slower, repeated, and short sentences for both voice-enabled personal assistants. They felt that Alexa can be difficult to hear and complicated and that Google Home was "*a very fast talker and needed to slow down*". With regards to asking Google Home or Alexa for cooking instructions, both end-users found Google Home to be better as Alexa offers too much information that is "*difficult to retain in one go*". They also found the instructions given by Alexa to be a bit more complicated than the Google Home ones and found it hard to remember the number of meal choices offered by Alexa. Similar to [32], both end-users thought that Google Home "*sounds good*" with the older woman preferring the female voice to the male one. Although they did not find the instructions to be quick, her daughter mentioned that it might be "*a bit fast whilst doing it but the pace is ok whilst listening*". Similar to the support workers taking part in the previous workshop, the older woman and her daughter were already using Alexa for medication reminders and were not aware of any additional functionalities the device offered such as asking for cooking recipes, playing music, turning off the lights in the entire house.

Certainty: The frail older man and his family members felt that IntraVox could provide independence, security and enhance all of their quality of life. Although they live together, he indicated that being able to alert his daughter when she was away could enhance his sense of security. She and her husband also suggested that this would provide "*peace of mind*" to them and would help enable them to do activities that they enjoyed.

Voice. The high level of personalization offered by IntraVox, in particular, received positive feedback from both stakeholder categories, with CoP members declaring: "*It is helpful to program systems to sound like a family member. This helps in personalising the support provided to the customer.*".

Silent Communication: Confirming previous end-users with LDA workshop and research findings [25, 57], the older woman stated that she would prefer voice instructions (regardless of voice) to silent ones (System 1 – Silent System).

Personification: Similar to the end-users with LDA, the frail older man and his family members provided positive feedback regarding the familiar voice that IntraVox provides and stated that the message delivered was at a good pace. They felt that "*IntraVox was the best technology from all the technologies* [systems] *presented*" but they also commented it might be suitable only for mild and moderate dementia (e.g., the familiar voice might confuse the user that the person is physically present in the room). The older woman also found the concept of a human voice enabling Alexa or Google Home to be interesting. She mentioned the computer voice (System 2 – Synthesized System) "*was alright*" but would prefer a human voice. She also found the verbal prompts (Hygiene and Dehydration Scenario) to be useful and would like a reminder to drink more liquids during the day. Her daughter also mentioned that it would be very useful from a carer's point of view to receive a text alert when certain events occur, such as setting up an alert when the door opens in the middle of the night (Reminder Scenario). The older woman did not have a problem sharing this data with her daughter mentioning that: "*I want you* (her daughter) *to know everything.*".

Adherence: CoP members stated that IntraVox "*could be phenomenal*" in helping people adhere to a routine, especially when they are encouraged to conduct tasks related to everyday life. The Activity Scenario received positive feedback and could be useful for people with low motivation.

Personalisation. CoP members highlighted that end-users can encounter difficulties in using technologies as a consequence of mood, lack of confidence, sensory and memory problems. Similar to [41] and as mentioned during previous workshops, CoP members explained that personalisation is a key priority and mentioned that IntraVox can ensure that end-users support would be individualized. End-users have very different requirements: "*Some are visual learners whereas others learn best through listening.*". Therefore, the technology offer should be grounded in a detailed assessment of the end-user's needs and aspirations. CoP members were also keen to stress that further development of IntraVox would be needed to ensure that keywords are easy to understand, and the content motivates end-users. Personalisation is a key aspect of assistive technologies, especially in situations where a person's cognitive abilities are constantly changing [34]. Therefore, CoP members argued that technology should be empowering as well as enabling the end-user. It is important that the end-user could override the automated processes (e.g., switching lighting on/off) or that the system could be altered when the end-users needs change.

5 Discussion

Our studies show how voice-enabled personal assistants can help improve the quality of life of people with complex needs. Similar to previous findings [38], we also encountered issues when engaging with such devices. Here, we discuss some opportunities to improve interaction and user experience.

IntraVox. People suffering from cognitive and speech impairments can encounter major problems in communicating with voice-enabled personal assistants. The need to vocalize the instruction can require a lot of effort for these user categories and can lead to frustration and irritation when the device does not comprehend the command or does not reply accordingly. In our studies, we deliberately used a recording of a human voice (this could be of a relative, a caregiver, a friend) to enable the voice-enabled personal assistant. For some users, this can create a sense of security and comfort, as if the person was there to help, rather than commands being carried out automatically by an anonymous computer or by using a synthetic voice. In our demonstrations, this use of a familiar, recognizable voice for Computer-Computer interaction has received positive feedback from end-users and social and healthcare organizations. IntraVox is also agile; as the purpose evolves, it can be changed over time in response to reflection and assessment of the end-users needs. This change in use could easily be implemented by simply changing the position of sensors and re-programming the message provided by IntraVox. This agile feature can be embedded in an individual's habitual practices and ensures responsiveness to ever-changing needs and health conditions.

We believe that the human voice used in our demonstrations can have an impact on the user experience [51, 48] and can play an important role in providing personalisation. Synthetic voices can affect people's social presence feeling of a device and customizing a computer voice according to the user's personality can rise the feeling of social presence [30]. Voice can also have a significant impact on trust and credibility [20, 29], with previous studies [38] showing that the emotion in the voice can influence an individual's perception of the content (more than a synthesized speech does). Our findings show that the person the voice belongs to and their relationship with the end-user is also important, as one end-user with LDA stated they would prefer the carer's voice and not the voice of a relative. The familiar, recognizable voice can also be beneficial in learning interventions [9, 36] and can help users who have difficulties verbalizing or remembering commands. Previous research [58] shows that older adults can benefit from listening to familiar voices as it can help them identify words in sentences. We believe that IntraVox could have the added benefit of reinforcing the learning of complicated verbal commands. This would allow users to take control of the situation as they learn the commands and take appropriate actions themselves instead of letting the system do it.

Things That Talk. "*Things That Talk*" has been used as a way of describing the Internet of Things (IoT), i.e. the notion that sensors and actuators in the physical world may communicate wirelessly and send data to each other [23]. The paradigm is related to the well-established concepts of Ubiquitous and Pervasive Computing. In our case, we are interested in the use of IoT in the home environment to control smart appliances. However, the way that IoT and home automation are implemented at present is in most cases hidden from human understanding – the things do talk to each other, but we cannot hear them! With IntraVox, by using an audible voice rather than "*silent*" communication between devices, we are purposefully aiming to reveal the inner workings of an IoT system. Rather than things happening without explanation when a sensor is triggered (similar to [57], one end-user described as the house being "*haunted*"), the system explicitly expresses its intentions by speaking out loud.

It is important to stress that this is not a way to increase efficiency – from a purely technical point of view, it makes more sense to send commands directly between systems using a silent communication protocol rather than using voice. Without considering any marketing perspectives [40], we believe that the human element that IntraVox introduces to home automation has a value in itself, something which was also confirmed in our evaluation workshops. In the field of Human-Robot Interaction, Tan et al. [53] show that people perceive robots as more likable if they see them humanly treating each other. Therefore, the use of a personalized human voice within the IntraVox system (rather than anonymous speech synthesis) could create a sense of security and comfort, rather than commands being carried out automatically by an unseen agent. Similar to previous Robot-to-Robot Communication studies in the presence of humans [25, 57], our participants appreciated being able to "*hear what will happen*" and know why a certain action would occur.

We think that some principles arising out of this can be applied in other situations, going beyond the users that IntraVox was originally developed for. One way to situate this is to relate to the debate on interface agents versus direct manipulation in HCI [49]. Agents act autonomously, often without the user's conscious input; direct manipulation, on the other hand, puts the control explicitly in the user's hands. In our experiments, the silent control system is akin to an agent, whereas giving commands directly to a voice assistant corresponds to direct manipulation. IntraVox, however, falls in between these 2 interface approaches, as on one hand, it relies on sensors to initiate events, but on the other, it allows the user to be aware of these actions and if needed insert themselves in the process (for instance by counter-acting a command that was issued). We believe IntraVox can suggest an interface principle that we call *Things That Talk (Among Themselves)* - IoT devices and other context-aware systems that make an event explicit and visible (or audible) to the user, rather than hidden.

6 Limitations and Future Work

IntraVox uses a pre-recorded human voice as the text-to-speech voice generated by the Raspberry Pi was abrasive and often misinterpreted. We started by using free text-to-speech Python libraries as this was the easiest way to produce audio files in terms of time and resources. Using more advanced voice synthesis, e.g. Deep Voice [3] it will be possible to generate not just a more pleasing voice, but also one that is reminiscent of a known and trustworthy person.

There is a need to widen our CoP members to include speech and language therapists to create IntraVox messages that are easy to understand and follow. Instructing the user about what words to say to trigger an action (e.g. "*To close the curtains, say, 'Alexa, close the curtains.'*") can empower them to make their own decisions. In the current implementation, users cannot reject and stop the command, unless they are able to reverse the command once it is being carried out. One way to overcome this is by using speech recognition and provide users with options: "*Do you want me to ask Alexa to close the curtains? Say yes or no.*". Taking the user's physical presence in the room should also be considered to avoid instances when the system announces its action while the user is out of the room. The resulting user experience of returning to a room to find the blinds closed could be perceived as "*haunted*" as the Silent system.

We also aim to implement a platform to ease the setup process, e.g. select the person the voice recording should belong to and record the audio files, select the desired commands and sensor values. It might also be that the system can learn to change the commands based on the user's behaviour. For example, the system could detect whether the users followed the suggestion of turning on the lights when it's getting dark outside and if not, the system can do it for them.

Alexa uses a built-in technology that is constantly listening for the wake word - the user has little knowledge of when Alexa might be actively listening to them and their environment. We do acknowledge that the microphone must always be enabled to carry out the commands. One way to overcome this aspect would be by controlling Alexa using third-party websites that do not require its microphone to be constantly enabled.

7 Conclusion

In this paper, we introduced IntraVox, a novel voice-based interaction system that allows a Computer-Computer Interaction using a highly personalized, audible human voice. The system was developed and evaluated with multiple stakeholder categories including older and frail people and people with LDA. The system introduces the concept of home automation by allowing sensors, actuators, and IoT devices to verbally communicate with each other using a highly personalized, human voice. This solution has the potential to reinforce the learning of specific commands and make visible the inner workings of a home automation system, removing anxiety and creating a sense of comfort. Situated in between agents and direct manipulation, this approach makes visible the often-hidden workings of a ubiquitous computing system. Our current example uses audible voice commands, but other modalities could also be explored, creating more user-friendly and understandable home automation and IoT environments. We believe that although the initial intended user groups are those with complex needs, many of the lessons learned can be generalize to the general population, making ubicomp systems more understandable, comforting, and easy to use.

References

1. Adib, F., Mao, H., Kabelac, Z., Katabi, D., Miller. R.: Smart homes that monitor breathing and heart rate. In Proc. of the 33rd annual ACM conference on human factors in computing systems. 2015.
2. Amazon Skills Kit, https://amzn.to/363uZUU. Last accessed 23 Jan 2021
3. Arik, S.- Ö., Chrzanowski, M., Coates, A., Diamos, G., Gibiansky, A., Kang,Y., Li, X., Miller, J., Ng, A., Raiman, J., Sengupta,S., Shoeybi, M.:Deep voice: real-time neural text-to-speech. In Proc. of the 34th International Conference on Machine Learning - Volume 70 (ICML'17). JMLR.org, 195–204. 2017.
4. Atefi, S.,Truelove, A., Rheinschmitt, M., Almeida, E., Ahmed, I., Alipour, A.:2020. Examining user reviews of conversational systems: a case study of Alexa skills. arXiv preprint arXiv:2003.00919. 2020.
5. Audacity, https://www.audacityteam.org/

6. Bakhai, A., Constantin, A., Alexandru. C. A.: Motivate Me! An Alexa Skill to Support Higher Education Students with Autism. International Conferences Interfaces and Human Computer Interaction and Game and Entertainment Technologies. 2020.
7. Barnes, J., Richie, E., Lin, Q., Jeon, M., Park, C.-H.: Emotive Voice Acceptance in Human-Robot Interaction. In Proc. of the 24th International Conference on Auditory Display. 2018.
8. Bespoken, https://bit.ly/2LO9rVp. Last accessed 23 Jan 2021
9. Brewer, R.-N., Morris, M.-R., Lindley, S.-E.: How to remember what to remember: exploring possibilities for digital reminder systems. In Proc. of the ACM on interactive, mobile, wearable and ubiquitous technologies 1.3 : 1-20. 2017.
10. Burgoon, J.-K., Bonito, J.-A., Bengtsson, B., Cederberg, C., Lundeberg, M., Allspach, L.: Interactivity in human–computer interaction: A study of credibility, under-standing, and influence. Computers in Human Behavior,16(6), 553–574. 2000.
11. Carroll, C., Chiodo, C., Lin, A.-X., Nidever, M., Prathipati, J.: Robin: enabling independence for individuals with cognitive disabilities using voice assistive technology. In Proc. of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (pp. 46-53). 2017.
12. Cho, M., Lee, S.-S., Lee, K.-P.: Once a Kind Friend is Now a Thing: Understanding How Conversational Agents at Home are Forgotten. In Proc. of the 2019 on Designing Interactive Systems Conference, pp. 1557-1569. 2019.
13. Colley, M.: Use of frequency volume charts and voiding diaries. Nursing Times, 111(5). 2015.
14. Corbett, E., Weber, A.: What can I say? addressing user experience challenges of a mobile voice user interface for accessibility. In Proc. of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (pp. 72-82). 2016.
15. Couper, M.-P., Singer, E., Tourangeau, R.: Does voice matter? An interactive voice response (IVR) experiment. Journal of official statistics 20, no. 3: 551. 2004.
16. Cruz-Sandoval, D., Favela, J.: Co-designing ambient-assisted interventions using digital interlocutors for people with dementia. In Proc. of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers, pp. 813-821.
17. Cruz-Sandoval, D., Penaloza, C.-I., Favela, J., Castro-Coronel, A.-P.:Towards social robots that support exercise therapies for persons with dementia. In Proc. of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, pp. 1729-1734.
18. Daher, M., El Najjar, M., Diab, A., Khalil, M., Charpillet, F.: Multi-sensory Assistive Living System for Elderly In-home Staying. In 2018 International Conference on Computer and Applications (ICCA), pp. 168-171. IEEE. 2018.
19. Dahmen, D., Minor, B., Cook, D., Vo,T., Schmitter-Edgecombe, M.: Smart home-driven digital memory notebook support of activity self-management for older adults. Gerontechnology 17: 113-125. .2018.
20. Geeng, C., Franziska Roesner, F.: Who's In Control? Interactions In Multi-User Smart Homes. In Proc. of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1-13). 2019.
21. Home Assistant, https://www.home-assistant.io/. Last accessed 23 Jan 2021
22. Hsieh, H.-F., Shannon, S.: Three approaches to qualitative content analysis. Qualitative health research 15.9: 1277-1288. 2005.
23. Igoe, T.: Making Things Talk: Practical Methods for Connecting Physical Objects, O Reilly; 3rd ed. 2017.

24. Juniper Research, https://bit.ly/3iGO8Ri. Last accessed 23 Jan 2021
25. Kanda, T., Ishiguro, H., Ono, T., Imai, M., Mase, K.: Multi-robot cooperation for human-robot communication. In Proc. 11th IEEE International Workshop on Robot and Human Interactive Communication, pp. 271-276. IEEE. 2002.
26. Kim, D., Park, K., Park, Y., Ju, J., Ahn, J.-H.:Alexa, Tell Me More: The Effect of Advertisements on Memory Accuracy from Smart Speakers. In PACIS. 2018.
27. Kiseleva, J., Williams, K., Jiang, J., Awadallah, A. H., Crook, A. C., Zitouni, I., Anastasakos, T.: Understanding user satisfaction with intelligent assistants. In Proc. of the 2016 ACM on Conference on Human Information Interaction and Retrieval, pp. 121-130. 2016.
28. Kowalski, J., Jaskulska, A., Skorupska, K., Abramczuk, K., Biele, C., Kopeć, W., Marasek, K.: Older Adults and Voice Interaction: A Pilot Study with Google Home. In Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. 2019.
29. Large, D. R., Burnett, G. E., The effect of different navigation voices on trust and attention while using in-vehicle navigation systems. Journal of safety research 49 : 69-e1. 2014.
30. Lee, K.-M., Nass, C.: Social-psychological origins of feelings of presence: Creating social presence with machine-generated voices. Media Psychology 7, no. 1. 2005.
31. Li, X., Zhang, D., Lv, Q., Xiong, J., Li, S., Zhang, Y., Mei, H.: IndoTrack: Device-free indoor human tracking with commodity Wi-Fi. In Proc. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1, no. 3: 1-22.2017.
32. Lombard, M., Ditton, T.: At the heart of it all: The concept of presence. Journal of computer-mediated communication 3, no. 2: JCMC321. 1997.
33. Luger, E., Sellen, A.: " Like Having a Really Bad PA" The Gulf between User Expectation and Experience of Conversational Agents. In Proc. of the 2016 CHI Conference on Human Factors in Computing Systems. 2016.
34. McGee-Lennon, M. R., Gray, P. D.: Including stakeholders in the design of home care systems: Identification and categorisation of complex user requirements. INCLUDE Conference. 2007.
35. Mennicken, S., Zihler, O., Juldaschewa, F., Molnar, V., Aggeler, D., Huang, E. M.:" It's like living with a friendly stranger" perceptions of personality traits in a smart home. In Proc. of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 120-131. 2016.
36. Milani, A., Lorusso, M. L., Molteni, M.: The effects of audiobooks on the psychosocial adjustment of pre-adolescents and adolescents with dyslexia. Dyslexia 16, no. 1: 87-97. 2010.
37. Nass, C. I., Brave, S.: Wired for speech: How voice activates and advances the human-computer relationship. Cambridge, MA: MIT press. 2005.
38. Nass, C. I., Foehr, U., Brave, S., Somoza, M.: The effects of emotion of voice in synthesized and recorded speech. In Proc. of the AAAI symposium emotional and intelligent II: The tangled knot of social cognition. North Falmouth. 2001.
39. Nass, C. I., Steuer, J., Tauber, E. R.: Computers are social actors. In Proc. of the SIGCHI conference on Human factors in computing systems, pp. 72-78. 1994.
40. Oppenheimer, A.: From experience: products talking to people—conversation closes the gap between products and consumers. Journal of Product Innovation Management 22, no. 1: 82-91. 2005.
41. Perelmutter, B., McGregor, K. K., Gordon, K. R.:Assistive technology interventions for adolescents and adults with learning disabilities: An evidence-based systematic review and meta-analysis. Computers and education 114: 139-163., 2017.

42. Plivo, https://www.plivo.com/. Last accessed 23 Jan 2021
43. Porcheron, M., Fischer, J. E., Reeves, S., Sharples, S.: Voice interfaces in everyday life. In proc. of the 2018 CHI conference on human factors in computing systems (pp. 1-12). 2018.
44. Pradhan, A., Mehta, K., Findlater, L.: "Accessibility Came by Accident" Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In Proc. of the 2018 CHI Conference on Human Factors in Computing Systems,pp. 1-13.
45. Purington, A., Taft, J. G., Sannon, S., Bazarova, N. N., Taylor, S. H.: " Alexa is my new BFF" Social Roles, User Satisfaction, and Personification of the Amazon Echo. In Proc. of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, pp. 2853-2859. 2017.
46. Rafferty, J., Nugent, C.-D., Liu, J., Chen, L.: From activity recognition to intention recognition for assisted living within smart homes. IEEE Transactions on Human-Machine Systems 47, no. 3. 2017.
47. RPi Text to Speech, https://bit.ly/3poEAxf, Last accessed 23 Jan 2021
48. Sayago, S., Neves, B. B., and Cowan, B. R.: Voice assistants and older people: some open issues. In Proceedings of the 1st International Conference on Conversational User Interfaces (pp. 1-3). 2019.
49. Shneiderman, B., Maes, P.: Direct manipulation vs. interface agents. Interactions 4, 6, 42–61. 1997.
50. Simpson, J., Gaiser, F., Macík, M., and Breßgott, T.: Daisy: A Friendly Conversational Agent for Older Adults. In Proceedings of the 2nd Conference on Conversational User Interfaces (pp. 1-3).2020.
51. Søndergaard, M. L. K., Hansen, L. K.: Intimate Futures: Staying with the Trouble of Digital Personal Assistants through Design Fiction. In Proceedings of the 2018 Designing Interactive Systems Conference, pp. 869-880. 2018.
52. Suzuki, Y., Kato, K., Naomi Furui, N., Sakamoto, D., Sugiura, Y.: Cushion Interface for Smart Home Control. In Proc. of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction (pp. 467-472). 2020.
53. Tan, X. Z., Reig, S., Carter, E. J., Steinfeld, A.: From one to another: how robot-robot interaction affects users' perceptions following a transition between robots. In 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 114-122. IEEE. 2019.
54. Twilio, https://www.twilio.com/, Last accessed 23 Jan 2021
55. Vacher, M., Caffiau, S., Portet, F., Meillon, B., Roux, C., Elias, E., Lecouteux, B., Chahuara, P.: Evaluation of a context-aware voice interface for ambient assisted living: qualitative user study vs. quantitative system evaluation. ACM Transactions on Accessible Computing (TACCESS) 7, no. 2: 1-36. 2015.
56. Vines, J., Lindsay, S., Pritchard, G.-W., Lie, M., Greathead, D., Olivier, P., Brittain, K.: Making family care work: dependence, privacy and remote home monitoring telecare systems. In Proc. of the 2013 ACM international joint conference on Pervasive and ubiquitous computing. 2013.
57. Williams, T., Briggs, P., Scheutz, M.: Covert robot-robot communication: Human perceptions and implications for human-robot interaction. Journal of Human-Robot Interaction 4.2: 24-49. 2015.
58. Yonan, C. A., Sommers, M. S.:The effects of talker familiarity on spoken word identification in younger and older listeners. Psychology and aging, 15.1: 88. 2000.