

Voice Assistants and Smart Speakers in Everyday Life and in Education

George TERZOPOULOS, Maya SATRATZEMI

Department of Applied Informatics, University of Macedonia, Thessaloniki, Greece
Email: gterzopoulos@uom.edu.gr, maya@uom.edu.gr

Received: November 2019

Abstract. In recent years, Artificial Intelligence (AI) has shown significant progress and its potential is growing. An application area of AI is Natural Language Processing (NLP). Voice assistants incorporate AI by using cloud computing and can communicate with the users in natural language. Voice assistants are easy to use and thus there are millions of devices that incorporate them in households nowadays. Most common devices with voice assistants are smart speakers and they have just started to be used in schools and universities. The purpose of this paper is to study how voice assistants and smart speakers are used in everyday life and whether there is potential in order for them to be used for educational purposes.

Keywords: artificial intelligence, smart speakers, voice assistants, education.

1. Introduction

Emerging technologies like virtual reality, augmented reality and voice interaction are reshaping the way people engage with the world and transforming digital experiences. Voice control is the next evolution of human-machine interaction, thanks to advances in cloud computing, Artificial Intelligence (AI) and the Internet of Things (IoT). In the last years, the heavy use of smartphones led to the appearance of voice assistants such as Apple's Siri, Google's Assistant, Microsoft's Cortana and Amazon's Alexa. Voice assistants use technologies like voice recognition, speech synthesis, and Natural Language Processing (NLP) to provide services to the users. A voice interface is essential for IoT devices that lack touch capabilities (Metz, 2014). Besides smartphones, voice assistants are now incorporated in devices that are equipped with a microphone and a speaker to communicate with the users, called smart speakers.

Cloud platforms are now enabling voice assistants in millions of homes. Voice assistants rely on a cloud-based architecture, since data has to be sent back and forth to centralized data centers. A smart speaker is relatively simple by design, which means most of the computing and artificial intelligence processing happens in the cloud and

not in the device itself. The basic idea is that the user makes a request through the voice-activated device, and then, the voice request gets streamed through the cloud, and here voice gets converted into text. Then, the text request goes to the backend and after processing, the backend replies with a text response. Finally, the text response goes through the cloud and gets transformed into voice, which will be streamed back to the user. Most smart speakers come without a screen although there are smart speakers with screens such as the Amazon Echo Show and Echo Spot, the Facebook Portal, and the Google Home Hub. The popularity of these devices is constantly rising since 2017. According to Canalys (2018), smart speaker installed base will approach 225 million by 2020 and 320 million by 2022. Amazon Echo and Google Home devices are considered to reside in over 50% of US households by 2022 and global ad-spending on voice assistants will reach \$19 billion by the same year according to Juniper Research (2017). The Alexa platform is the dominant market leader, with more than 70% of all intelligent voice assistant-enabled devices (other than phones), running the Alexa platform (Griswold, 2018).

Voice assistants have several interesting capabilities such as:

- Answer to questions asked by users.
- Play music from streaming music services.
- Set timers or alarms.
- Play games.
- Make calls or send messages.
- Make purchases.
- Provide information about the weather.
- Control other smart devices (lights, locks, thermostats, vacuum cleaners, switches).

The capabilities of voice assistants are continuously extending. Amazon and Google have provided platforms for developers in order to extend their assistants' capabilities. Similar to mobile apps, Amazon Skills and Google Actions, radically expand assistants' repertoire, allowing users to perform more actions with voice-activated control.

According to Sheppard (2017), some key elements that distinguish voice assistants from ordinary programs are:

- NLP: the ability to understand and process human languages. It is important in order to fill the gap in communication between humans and machines
- The ability to use stored information and data and use it to draw new conclusions
- Machine learning: the ability to adapt to new things by identifying patterns

Similarities and differences of devices and services regarding voice assistants have been studied in the literature (López *et al.*, 2017; Kępuska and Bohouta, 2018). In addition, as with any new revolutionary technology, scientific research and the educational community are considering whether these new devices can help the educational process. Something similar has happened before with personal computers and tablets (Algoufi, 2016; Gikas and Grant, 2013; Herrington and Herrington, 2007).

The purpose of our paper is to present findings regarding home usage of voice assistants and smart speakers, as well as some early attempts for using them for educational

purposes. Although voice assistants are present in many homes, their use in school environments and for educational purposes is limited since there are many concerns regarding their privacy settings and data collection. Study of home usage will provide insights regarding the ease of use of this new technology and how users perceive it. Furthermore, education can take place in formal or informal settings, thus it is evident to examine the use of voice assistants and smart speakers, inside or outside the classroom and by children, adults and elderly people.

Our specific research questions (RQ) for the study are as follows:

- **RQ 1:** How do children, adults and elderly people use voice assistants and smart speakers in their everyday life?
- **RQ 2:** How have voice assistant and smart speaker technologies been used in education?
- **RQ3:** What kind of security concerns do users have, regarding the use of voice assistants and smart speakers?

The remainder of the paper is organized as follows. In Section 2, the methodology to retrieve related papers is described, while in Section 3, studies about smart speakers' home usage by people of every age are presented. Section 4 includes related work about AI, voice assistants and smart speakers uses for educational purposes. The educational process can concern small children (kindergarten), children (primary education), teenagers (secondary education), adults and elderly people (lifelong learning). It also includes people with disabilities (special education). Section 5 raises the security and privacy concerns pointed out by many researchers and users. Since privacy is a major issue, it is evident that for voice assistants to be used in a classroom setting and for educational purposes, all security issues should be resolved. Finally, Section 6 interprets the findings of this study while in Section 7, new areas for future research are recommended.

2. Methodology

In order to retrieve sufficient and high-quality papers regarding uses of voice assistants and smart speakers, the snowball technique as described by Wohlin (2014) was used. The technique has the following steps:

- Initially perform a search in Google Scholar, IEEE Xplore, Scopus and ACM Digital Library and gather the initial start set of relevant papers. Keywords used were “voice assistant”, “smart speaker”, “amazon echo”, “google assistant”, “Alexa” and “Siri”.
- For the initial start set of papers, iterate through backward and forward snowballing. Backward snowballing uses the reference list to identify new papers to include, while forward snowballing refers to identifying new papers that cite the paper being examined. With backward and forward snowballing, new papers that are identified in each step, are put into a pile to go into the next iteration.

By using the snowball technique, 37 scientific papers were retrieved, all of them presented in this study.

3. Home usage

Adults: There are few studies that explore the usage of smart speakers in homes and users' satisfaction. Bunyard (2019) provides insights regarding the complex reasons that people adopt Internet of Things technology into their lives. The main reason is the convenience that the technology offers since users don't have to deal with things that take time and cause stress. Purington *et al.* (2017), explored the degree of personification of the Amazon Echo devices, the sociability level of interactions and users' satisfaction, based on a total of 851 user reviews of the Amazon Echo, posted on Amazon.com. Results indicate that there are variations in how people refer to the technology, with over half using the personified name "Alexa", and there is a moderate degree of sociability. Users report that they interact with the device for entertainment purposes such as listening to music, or for other functions like retrieving information, manage scheduling and shopping.

Interesting findings came from Sciuto *et al.* (2018), where authors explored how households incorporate conversational agents into their lives. Specifically, authors analyzed the logs of 75 Alexa users, for a total of 278,654 voice commands. Participants who have owned an Alexa device for at least six months, answered survey questions related to their household use of Alexa. Of the 75 participants, 26 reported having children although data from the log files did not provide any insights into which household member gave each command. Parents that were interviewed, positively recalled their children successfully interacting with Alexa even before interacting with smartphones and other technology devices.

Data from 724 participants using Amazon Echo in the UK were gathered by McLean and Osei-Frimpong (2019). Participants had used the device for at least one month to provide insight into the variables motivating the use of the in-home voice assistant. Information was collected using questionnaires. Findings revealed that voice assistants are used for utilitarian purposes, in order to help people complete tasks, look up information, seek support and process orders.

Furthermore, by interviewing 31 participants, Rzepka (2019) analyzed the benefits and costs that users evaluate when using voice assistants. The study concludes that the fundamental objectives that maximize users' overall value of using voice assistants, are efficiency, convenience, ease of use, minimal cognitive effort, and enjoyment. Voice assistants can be operated without the use of the hands and without thinking about syntax or grammar errors compared to using text as input. Participants mentioned that they enjoyed interaction and were curious about the answers they were provided.

In a study by Song (2019), 433 adult participants completed an online survey, in order to assess perceived usefulness, perceived ease of use, attitude towards voice assistants, and behavioral intention to use them. Findings suggest that perceived usefulness has significant effects on individuals' attitude toward voice assistants and behavioral intention to adopt this technology. Furthermore, consumers seek to buy devices that are easy to use.

Children: Some studies targeted children behavior towards voice assistants, in order to assess how children interact with them, what are they using them for and whether they are having trouble in communicating with them.

Beirl *et al.* (2019), conducted a research about the home usage of Alexa, in a period of three weeks. The purpose of the study was to investigate how families learn the new Alexa skills regarding music, storytelling and games and appropriate them into their lives. In order to collect the data regarding how and when the skills were used, researchers collected voice recordings and conducted interviews. Six families with children in the age group of 2–13 years were recruited. Results showed that there was much enthusiasm about how they had interacted with Alexa and how it became part of their family rituals. The interactions with Alexa often resulted in much shared laughter and there were also several instances of teasing. There was also a lot of encouragement, specifically when a more competent family member helped a younger member interact with Alexa. The study concluded that all of the above interactions, contributed to social and emotional bonding, leading to further family cohesion. Another important finding of the study was that when younger children were having trouble following the rules of playing a game or a quiz, families adopted helper roles to encourage and make suggestions for what younger children should say to Alexa.

Children behavior is investigated by Druga *et al.* (2017) where 26 participants (3–10 years old) interacted with 4 voice assistants, Amazon Alexa, Google Home, Cozmo, and Julie Chatbot. Children were divided into groups of 4–5 and played with each voice assistant for 15 minutes. After each session with a voice assistant, children answered a questionnaire, in the form of a game, in order to analyze children's perception of the voice assistant. Authors also interviewed 5 children, to further probe their reasoning. Children enjoyed interaction with voice assistants, while older children perceived their intelligence and thought they could learn from them. The main issue of the interaction with children was getting the assistants to understand their questions although with the help of facilitators and parents, children altered their strategy and became fluent in voice interaction.

Yuan *et al.* (2019), observed 87 children with ages 5–12 and 27 adults interacting with three Wizard-of-Oz speech interfaces. Children participants were recruited along with a parent or guardian who could provide consent, and potentially participate in the study. Answers with the Wizard-of-Oz technique were provided by humans although they were spoken by a computer program. Nevertheless, none of the children expressed suspicion or inquired about how the system worked. After reviewing the logs and audio recordings of all the participants, authors came to the conclusion, that children preferred personified interfaces rather than non-personified and that age played an important role in children's performance. Older children could get the answer that they needed using less help from provided hints. Since the interaction required children to reformulate questions, most of them needed hints to complete the task. Another interesting finding from this study, was that 93% of children had used one or more speech interfaces prior to the study and most children used such interfaces multiple times per day.

Children aged 5 to 6 and their parents' interactions with a smart speaker were also studied by Lovato *et al.* (2019). The study lasted two weeks and involved 18 families.

Each parent and child pair were interviewed in order to capture their view of the experience using the smart speaker. Analysis of the results showed that 89% of children's questions were transcribed correctly, although only 50% of children's questions received a full answer. Children and their parents reported that the provided answers were long or required interpretation. Most children's questions were about the world around them and they believed that the device is a source of information.

In a six-month study by Nilsen and Røynealand (2019) with preschool children 4–6 years old in Norway, authors came to the conclusion that the interaction with voice assistants is very fragile when young children are involved. This is because young children have not learned yet to master their conversational abilities and they are often playful and impulsive. Furthermore, it was observed that the children applied a make-it-up-while-you-go approach, and they did not always seem to have a specific purpose when they addressed questions to the voice assistant. Children were seeing the voice assistant as a social partner and wanted to know more about the personality of the voice assistant. They were also telling stories about themselves. Since all of the children had never interacted before with a voice assistant, it proved difficult for them to understand that the voice assistant needed to be activated first before they could talk to it. Finally, there were many times when an adult had to intervene and to support children during their interactions with the voice assistant, since children were often distracted and did not respond correctly to questions by the voice assistant.

It is evident that smart speakers have the potential to play an important role in children's self-directed learning. As Danovitch and Alzahabi (2013) showed, children as young as 5 years old, take past experience into account when choosing sources of information, both human and electronic, thus children who experience difficulty interacting with voice assistants, might be reluctant to use them again in the future.

Elderly people: The use of voice assistants by older people (65+) is investigated by Savago *et al.* (2019). As most digital technologies are designed by and for young people, without considering older populations, authors address the need to further research the use of this technology by older people. Kowalski *et al.* (2019) also investigated the use of voice assistants by older adults. Seven older adults were the participants of the study. All of them lived in Poland and their English language proficiency varied. The younger participant was 64 years old while the older was 89. Participants were described as active users with basic computer skills. Findings of the study suggest that the range of possibilities of voice assistants impressed the participants. Participants also liked the idea of being able to accomplish certain tasks using only speech not only because they might have difficulties moving as one participant mentioned, but also because it does not disturb their process. It must be noted that voice assistants were used as search engines mainly, but more importantly as a translator or even as a teacher. Authors also point out the need for further research regarding the use of voice assistants by older people.

People with disabilities: Few studies targeted users with cognitive disabilities or vision problems, in order to investigate how voice interfaces can assist them in their everyday life and how easy is for them to use voice assistants.

Baldauf *et al.* (2018), studied speech-based conversational interfaces for the cognitively impaired, i.e. people with neurological disorders with minor impairment in

instrumental activities of daily living including reading and writing difficulties. They also note that applications of conversational interfaces and voice assistants for people with cognitive impairments are scarce. The participants of the study emphasized the interest and motivation to use novel technical solutions such as voice assistants and have high expectations regarding conversational user interfaces. They believe that voice assistants should not be passive but to initiate a conversation with the users in certain conditions. They also point out that voice assistants should be a complement and not a replacement to personal contact with other humans. Authors conclude that voice assistants can act as a conversational partner and help to diminish the feeling of loneliness.

The popularity of smart speakers among people with disabilities is significant. By analyzing the reviews of verified purchases of Amazon Echo devices Pradhan *et al.* (2018) found out that almost 38% of reviews mentioned disabilities related to individuals with visual impairments or blindness, suggesting that voice assistants can be of great use by this community. Findings from Abdolrahmani *et al.* (2018) also confirm that voice interaction is convenient for blind people and help them make possible day-to-day tasks which other people may take for granted.

Vtyurina *et al.* (2019) conducted a survey with 53 people who are legally blind to identify the strengths and weaknesses of screen readers and voice assistants. Screen readers transform the visual content in a graphical user interface into audio by vocalizing on-screen text, thus they are an important accessibility tool for blind computer users. Voice assistants offer an alternative audio-based interaction paradigm. Participants reported that they were using both screen readers and voice assistants in their daily routine, multiple times. They also reported that voice assistants provide a direct answer with minimal effort but with limited insight, while screen readers and search engines provide the opportunity to review a number of different resources. With screen-readers, participants are able to customize multiple settings like speech rate and pitch to match their preferences while this ability is not provided by voice assistants. Additionally, voice assistants provide audio-based content that is not affected by the poor website design on some webpages. To combine the new technologies, authors developed VERSE (Voice Exploration, Retrieval and SEarch). VERSE was evaluated by 12 blind screen reader users who reported that the system was easy to learn and use, and expressed interest in its continued development or public release.

4. Education

Primary and Secondary education: In general, AI can be used in education in order to assist students in their learning process by providing fast and efficient information. For this purpose, Scarlet an Artificial Teaching Assistant is proposed by Ilhan *et al.* (2017), although it is not tested in an educational environment yet. Scarlet is comprised by 3 modules: natural language processing, pseudo contextual data analysis and trial and error learning. Mulyana and Hakimi (2018) proposed an AI chatbot-based virtual assistant called LTKA-Bot. It is a virtual assistant to provide services regarding course activities

such as record session data and services related to attendance, task assignment and scoring management.

Trivedi (2018) proposed ProblemPal, an Alexa Skill that enables teachers to automatically generate practice content with voice commands. The skill can create practice questions about any topic by using APIs from Wikipedia, Wolfram Alpha, and Khan Academy. Questions generated, are then uploaded and shared with students on Google Classroom. Although the skill was not tested in class, the study claims that it can reduce the workload of teachers and can also be a valuable tool for students.

A classroom setting is proposed by Horn (2018). The author suggests that each classroom can contain enough microphones in order to recognize each student's voice and provide through voice assistants, personalized answers to each student's headphones. Alternatively, each classroom can have a question station where each student can go and ask a question to a smart speaker. Teachers should get real-time data from voice assistant so that they can intervene when they see fit. The devices used in the setting are not seen as replacements for teachers, rather than amplifiers for their work.

Results of the University of Idaho (UI) Echo Project that took place during the 2017–2018 academic year, are discussed by Dousay and Hall (2018). The initiative of the project was to investigate perceptions and challenges related to integrating AI in classrooms. The project impacted four school districts and 900 students with 90 Amazon Echo Dot devices. The Echo Dot device was selected due to its low cost and compatibility issues. Different types of people used the device, such as a music teacher, a math teacher, a school counselor, the administrator, the elementary teacher and the classroom assistant. Teachers were depending on Alexa for simplifying some classroom processes, especially with the use of timers and reminders. Students preferred to ask the voice assistant for information, rather looking up information on a computer or a tablet. Both parties were excited to use the device in the classroom and tried to find ways to use it in particular projects and situations and many of them were already using a similar device at home. The study also suggests the cooperation of teachers, professional programmers and computer science faculty, to create custom skills that were not found in the existing Alexa Skills catalog. The study did not include observational data or log files from the devices, but relied on self-report perception surveys and interviews.

The differences in student engagement when a teacher implements purposeful instruction on using the intelligent voice assistant Siri in upper elementary and middle school science classrooms is explored by Neiffer (2018). Student engagement is associated with student graduation rates. High student engagement leads to increased teacher's satisfaction and enjoyment. Findings of the study shows that the relationship between technology and learning is too complex to make broad assumptions and there is no clear association between integration of Siri in 5th grade/middle school science classrooms and increase in student engagement. Davie and Hilber (2018) created a quiz custom Alexa Skill about Scotland and used it with students ahead of an excursion. Students used the Amazon Echo device and found the skill to be entertaining.

Kloos *et al.* (2019) created Java-PAL using Google Actions, an application with which students learn about the basic concepts of Java programming. Authors describe

their decisions that have guided the design of Java-PAL based on Voice User Interface (VUI) guidelines. These guidelines include the maximization of relevant information provided at each interaction and the clear communication and cooperation with the user. Java-PAL was not evaluated, thus its education value is undetermined.

Since voice assistants are very good at math and in spelling words correctly, they can be used by primary school students for grammatical control and for verifying their results in math. In a study by Selak (2017) with elementary students in first and second grade, it was observed that students consistently confirmed their mathematical results and did not ask for help from their teacher.

Special education: Porayska-Pomsta *et al.* (2018) explores the use of AI in special education. The study examined the education efficacy of a learning environment called ECHOES and included 29 children with Autism Spectrum Conditions (ASC) aged 4–14 years old. ECHOES includes a virtual AI character called Andy which serves as a social partner for children with ASC. Children interacted with ANDY and a human practitioner, blending human and AI interaction. Results showed a significant increase in the proportion of children's responses to the human social partners and suggested positive trends with respect to children's initiations to both social partners.

Language learning: It is evident that smartphones and mobile learning have enabled more people to learn and to succeed in learning, wherever they are. Informal learning is a natural accompaniment to everyday life both inside and outside the classroom. For many people, language learning is a common lifelong pursuit in order to learn a new language or to improve existing knowledge. This skill is pursued widely outside the classroom on online platforms, like FluentU and Duolingo. Language learning is also an important topic covered in all schools and since the majority of students are taking foreign language learning classes, voice assistants can be used for this purpose. In any case, one of the biggest challenges for many language learners is finding opportunities to speak the language. Voice assistants now have the ability to speak many different languages and can be considered as a language-learning tool. Amazon Alexa speaks 7 languages, Google Assistant 13 and Siri 21. However, the number of available spoken languages by the assistants is constantly rising. Thus, voice assistants can be used as a language partner or to train users' pronunciation.

Dilon (2018), introduced a smart speaker with Alexa, initially to first and second-year students. Introductory questions proved to be a fun activity for students. Afterwards, some of the skills from the Alexa skills store were used. Students firstly used a pronunciation skill that required them to listen and repeat sentences. Speaking speed was proven to be too fast for them. A guessing game and an interactive story was used after, but the vocabulary levels were not well matched with learners' ability, and again the speaking speed was too fast. Authors used the Alexa Skills Kit to create a custom skill, an interactive game from a textbook page so that the students could self-study. By using the custom skill, students quickly identified where they were making pronunciation mistakes or providing the wrong answers. The study concludes that using voice assistants, students in different year groups were able to engage with each other in English and build connections between the content and goals of their various English courses.

An Alexa skill called “Japanese Flashcards” was developed by Skidmore and Moore (2019). Flashcards are commonly used to study vocabulary. The concept is that a learner is presented with one side of the flashcard and tries to recall what is on the other side. Although the skill was not evaluated, authors suggest that Alexa could provide valuable learning paradigms such as conversational role play and pronunciation training and can contribute to the emerging field of “voice-assisted language learning”.

Bekmyrza (2019) studied 5 chatbots created specifically for English language learners. Chatbots simulate a person’s speech. Chatbots are considered a convenient way that will not judge users strictly when they make mistakes and they can be used for personal training at any time during the day since they are available 24/7. In the study, chatbots were not evaluated by students. In live communication, people are also using gestures and facial expressions. With voice assistants, users must try to use the language in a right way so that they will understand the meaning.

5. Security and Privacy Concerns

Although the popularity of voice assistants is very high, the potential of them being used in a classroom environment or for educational purposes is not fully considered. This is mainly due to the fact that there are security and privacy concerns, as expressed in several studies (Pfeifle, 2018; Lei *et al.*, 2017), since devices must be listening at all times so that they can respond to users. Lau *et al.* (2018a), interviewed 17 smart speaker users and 17 non-users to find out their arguments for and against adopting this new technology and their privacy perceptions and concerns. Many non-users believe that these devices are not useful at all and companies are not to be trusted. On the other hand, smart speaker users have fewer privacy concerns and rely on companies to safeguard their personal data which think are not interesting to others. They also place the devices in the house based on accessibility and sound. Some parents reported that their children enjoyed using voice assistants. Most of them purchased the devices in order to be early adopters of the technology and among the first to use it. As a proposal, authors suggest different accounts for children and the ability to mute the devices with audio commands. They also suggest that an incognito mode, similar to web browsers, should be present in order to prevent data collection from the companies. To address privacy concerns for recording devices, McReynolds *et al.* (2017) propose indicators while recording.

Users’ perception of privacy and the use of privacy controls is studied by Lau *et al.* (2018b). Users reported that although they are aware of the ability to erase the audio logs of Google and Amazon devices or use the physical mute button, they did not use these functions. Users also reported that in some cases they use private browsing mode when surfing through the internet and they would like this to also be available in the case of smart speakers.

Privacy is also considered by Hoy (2018) since anyone with access to a voice-activated device can ask it questions, gather information about the accounts and services associated with the device, and ask it to perform tasks. As stated by Horn (2018), since

devices can distinct children's voices, their specific learning needs are certain to raise questions with the Children's Online Privacy Protection Act (COPPA) and the Family Educational Rights and Privacy Act (FERPA).

In a survey by Malkin *et al.* (2019) regarding 116 smart speaker owners' beliefs, attitudes, and concerns about the recordings that are made and shared by their devices (Amazon and Google), it was found that 56% of them did not know that their recordings were being permanently stored and that they could review them. It was observed, that while participants did not consider their own recordings important or sensitive, they were more protective of others' recordings, such as children and guests, and were strongly opposed to the use of their data by third parties. Only 3% of the participants review their recordings and deleted them. Additionally, only 5% of the participants used the mute button on their device while only 4% unplugged their device in order to stop listening. The survey concludes that privacy controls are underutilized.

The modality effects of voice assistants in the context of health information acquisition is studied by Cho (2019). Fifty-three undergraduate students from an undergraduate course at a southeastern university used a mobile phone with Google Assistant and a Google Home device. Participants were instructed to ask two types of health questions (i.e., less versus more sensitive health questions) by using voice or text as input. Results showed that voice versus text interactions significantly enhanced perceived social presence toward the voice assistant, but only when low sensitivity information was involved and when the users reported less privacy concerns. In contrast, when the information sensitivity and individual's privacy concerns were high, text was as impactful as voice to induce human-like perceptions of the voice assistant. Additionally, findings suggest that voice interactions elevated the feelings of having a social conversation between the user and the voice assistant and this led to positive evaluation toward the voice assistant. However, this occurred only when information being asked was less sensitive and while individuals reported low levels of privacy concerns.

6. Discussion

In the previous sections, research about the usage of voice assistants and smart speakers in homes and for educational purposes were presented. Studies regarding security and privacy concerns were also included.

Regarding RQ1, early findings from a small number of studies, show that adults (Purington *et al.*, 2017; Sciuto *et al.*, 2018; McLean and Osei-Frimpong, 2019; Rzepka, 2019; Song, 2019; Bunyard, 2019) use voice assistants for entertainment purposes, seeking information, making purchases and listening to music. Adults also enjoy interaction and find voice assistants easy to use. Studies involving children (Sciuto *et al.*, 2018; Beirl *et al.*, 2019; Druga *et al.*, 2017; Yuan *et al.*, 2019; Lovato *et al.*, 2019) conclude that children can interact with smart speakers and voice assistants, making basic requests. Children show much enthusiasm, believe that voice assistants are a source of information and can alter their strategy regarding the way they ask questions, becoming fluent in voice interaction.

Younger children near the age of 5 are having some troubles regarding voice interaction, although with a little help from their parents and their older siblings, they seem to manage. Children see voice assistants as a social partner, and they ask questions about them and sharing their stories. Additionally, children were found to be attracted and intrigued by the new technology, although as noted by Wiederhold (2018), it is vital for parents to teach their children that smart speakers and voice assistants are just tools and they should not rely heavily on technology during their daily routines.

Few studies with older people (Savago *et al.*, 2019; Kowalski *et al.*, 2019) indicate that they are impressed by this new technology and voice assistants not only can be used as a search engine, but also as a teacher and a translator.

Results from four related studies (Baldauf *et al.*, 2018; Pradhan *et al.*, 2018; Abdolrahmani *et al.*, 2018; Vtyurina *et al.*, 2019) show that voice assistants and smart speakers are already being used by people with cognitive disabilities or vision problems, although initially they were not designed for them. For the cognitively impaired, i.e. people with neurological disorders with minor impairment in instrumental activities of daily living including reading and writing difficulties, voice assistants not only can assist in their daily routine, but they can also act as a conversational partner (Baldauf *et al.*, 2018). Regarding blind people, Abdolrahmani *et al.* (2018) conclude that voice interaction is convenient for them and can play an important role in evoking feelings of independence and empowerment among them. For people with a broad range of disabilities, even some with hearing loss and speech impairments, Pradhan *et al.* (2018) found that by using voice assistants, they complete more independently everyday tasks. Compared to screen readers, Vtyurina *et al.* (2019) concluded that for blind people, voice assistants provide a direct answer with minimal effort but with limited insight, thus further improvements need to be made.

In conclusion, there is a variety of usages of voice assistants by each age group. This information gives great insight and some early indications whether voice assistants can be used and for other purposes, besides home entertainment and performing basic tasks.

Regarding RQ2, attempts on custom built AI assistants like Scarlet (Ilhan, 2017) and LTKA-Bot (Mulyana and Rifqy, 2018) and their usage in education are still in an early phase since Scarlet is not yet tested and LTKA-Bot is still under active development. The same applies for ProblemPal (Trivedi, 2018), an Alexa Skill that allows teachers to create practice content for their students for virtually any topic, since it is not tested yet.

Horn (2018) proposes a voice-activated classroom setting that will engage students, although the setting is not yet tested. It is believed that students from now on will be expecting individualized resources that smart speakers and voice assistants can provide, and not a passive environment.

A significant large-scale study by Dousay and Hall (2018) regarding 900 elementary students with 90 Amazon Echo Dot devices concluded that the usage of voice assistants and smart speakers in the classroom can be beneficial and exciting in most cases for both teachers and students. It must be noted that detailed data regarding students' age

were not presented and findings relied on self-report rather on observation. Furthermore, inconclusive findings come from Neiffer (2018) where authors claim that there is no clear association between integration of a voice assistant in 5th grade/middle school science classrooms and increase in student engagement.

At the university level, Davie and Hilber (2018) created a custom skill for Alexa Echo devices and all students found the device and the skill to be entertaining, although the study does not provide any input on sample size or any other result. Inconclusive input also applies in the study of Selak (2017), where echo devices were introduced to elementary students in first and second grade. There is not any detailed information regarding the sample size and the way data was collected. Additionally, the skill Java-PAL developed by Kloos *et al.* (2019) was also not evaluated. Additionally, the potential to facilitate autistic children's ability to engage in social interaction by using AI agents is demonstrated by Porayska-Pomsta *et al.* (2018).

Limited studies show that voice assistants and smart speakers can be used as an innovative tool for language learning. By using a custom Alexa skill with first and second-year students, Dillon (2018) concluded that students were able to engage with each other in English and build connections. The vocabulary custom skill, "Japanese Flashcards", created by Skidmore and Moore (2019) was not evaluated in class. Finally, 5 chatbots created specifically for English language learners were used in the study of Bekmyrza (2019) but they were also not evaluated by students. In conclusion, there are some early attempts regarding the use of voice assistants and smart speakers as language learning tools, although research in this topic is at its infancy.

In conclusion, regarding RQ2 and the use of voice assistant and smart speaker technologies in education, limited studies showed encouraging results even from early stages of education, i.e. kids at first and second grade of elementary school, although it is evident that more research on the topic is needed. Furthermore, children seem to be entertained by the use of the technology and this reflects to increased engagement. An important issue that has not been studied yet is the combined usage of voice assistants in homes and in the classroom. Since the popularity of voice assistants and smart speakers is rising, it would be interesting to see the devices as study companions in and out of the classroom. Early attempts in special education show that this new technology has some potential, although additional studies must examine the specialized needs of people with disabilities. Voice assistants can also play an important role regarding language learning although all languages are not supported yet. For students that their language is not supported yet, voice assistants can be used for learning foreign languages as well.

Regarding RQ3, seven related studies (Pfeifle, 2018; Lei *et al.*, 2017; Lau *et al.*, 2018a; McReynolds, 2017; Lau *et al.* 2018b, Malkin *et al.*, 2019; Cho, 2019) suggest that security and privacy concerns are major drawbacks for users of voice assistants and smart speakers. Some studies (Lau *et al.* 2018a; Lau *et al.*, 2018b) suggest that devices must be able to turn off by using voice commands instead of the physical mute button and that there should be an incognito mode where users can use the devices without be-

ing monitored. Other studies (McReynolds, 2017; Pfeifle, 2018), propose that devices should have obvious visual indicators regarding collection and transmission of data. Moreover, Lei *et al.* (2017) studied the security vulnerabilities of Amazon Echo Dot devices and discovered that acoustic attacks can be performed while victims are not at home. To sum up, there are a lot of concerns regarding privacy and security, although many people don't really care about these matters. Smart speaker companies should provide a more transparent view of the way that data is being used. As a result, in case of using the technology in schools, it is crucial that these concerns must be addressed, and all privacy issues are resolved.

Although there are some studies that suggest the use of voice assistants and smart speakers in a classroom setting, it is interesting to investigate whether this will assist the teaching process and achieve better pedagogical results. It is interesting that data from Organization for Economic Cooperation and Development (OECD, 2015) show that although the usage of Information and Communication Technology (ICT) in schools is rising, literacy levels and mathematical skills of students who have computers and other ICT equipment in their classrooms dropped. This issue is also addressed by Tsironis (2018), and the main reason that technology has not been as much helpful as expected to the education system comes from the lack of proper training for both students and teachers. Thus, the use of technology by itself is not enough. Proper training for teachers is needed in order to adopt and use smart speakers in the classroom. Additionally, as stated by Jahan (2017), AI can be helpful for learners providing personalized learning since different students respond differently to distinct motivations. Smart speakers should provide answers based on who is asking them and they can also include the name of who is asking in their response.

As noted by Kukulaska-Hulme (2019), there are important issues and developments in the field of "intelligent assistance", that increasingly encroach on the territory of human-to-human interaction and the idea of a friendly companion in the process of learning. There are also ethical concerns regarding deception, i.e. humans using technology to deceive other humans, or intelligent technology deceiving humans. Voice assistants and smart speakers could misunderstand learners or offer inappropriate advice. Since the availability of increasingly intelligent assistants on smartphones or smart speakers that are used in everyday life is rising, authors encourage the mobile learning research community to discuss these developments and their implications for research and practice, to shape a more extensive yet also targeted agenda for future work.

Since there are many concerns regarding privacy and security, smart speaker companies should address them in order for the devices to be used in the classroom. Studies show that smart speakers can be used as a source of knowledge in classrooms as they are now used in homes. Children can direct their questions to the devices and verify their results, instead of asking the teacher. Students are pushed to perform their questions in a right and comprehensible way by the devices, understanding the basic principles of human-computer interaction. Besides information, voice assistants and smart speakers can activate students and push them to search for knowledge.

7. Conclusion

Immersive learning technologies have the ability to update the existing education system. Virtual reality, augmented reality and voice assistants can provide new learning experiences. In this paper, research regarding the integration of AI voice assistants in education is presented. Research on this topic is limited since voice assistants and smart speakers are now gaining popularity. Findings presented in this paper will hopefully inspire other researchers to further investigate this topic. Smart speakers and voice assistants will be at the centre of interest in coming years as they enter the everyday life of households. The ways they can be used efficiently in the learning process is the subject of research as there are many challenges. One of these challenges is the lack of many languages as voice assistants do not speak all languages. In addition, voice assistants do not have many of the appropriate security measures and protection filters that can be used in class by students. Teachers need to be trained and motivated about the usefulness of these devices in order to adopt them in their class. Although in most cases positive results have been reported regarding students and teachers, results are limited, incomplete and unorganized. As a conclusion, the role of these devices and their use in the classroom are still at an early stage of research and more studies need to address this topic.

References

- Abdolahmani, A., Kuber, R., Branham, S. M. (2018). Siri Talks at You: An Empirical Investigation of Voice-Activated Personal Assistant (VAPA) Usage by Individuals Who Are Blind. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 249–258). ACM. DOI: 10.1145/3234695.3236344
- Algoufi, R. (2016). Using Tablet on Education. *World Journal of Education*, 6(3), 113–119.
- Baldauf, M., Bösch, R., Frei, C., Hautle, F., Jenny, M. (2018). Exploring requirements and opportunities of conversational user interfaces for the cognitively impaired. In *Proceedings of the 20th International Conference on human-computer interaction with mobile devices and services adjunct* (pp. 119–126). ACM. DOI: 10.1145/3236112.3236128
- Beirl, D., Yuill, N., & Rogers, Y. (2019). Using Voice Assistant Skills in Family Life. In Lund, K., Nicolai, G. P., Lavoué, E., Gweon, C. H., & Baker, M. (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019*, Volume 1 (pp. 96–103). Lyon, France: International Society of the Learning Sciences. DOI: 10.22318/csc2019.96
- Bekmyrza, K. (2019). *Using Chatbot to Increase Students' engagement in Education Process at High School*. [https://www.internauka.org/archive2/vestnik/11\(61_2\).pdf#page=68](https://www.internauka.org/archive2/vestnik/11(61_2).pdf#page=68)
- Bunyard, S. (2019). *Assistance from Alexa: The social and material benefits of the Internet of Things*. https://scholarcommons.scu.edu/cgi/viewcontent.cgi?article=1044&context=engl_176
- Canalys. (2018). *Media alert: Smart Speaker Installed Base to Hit 100 Million by End of 2018*. https://www.canalys.com/static/press_release/2018/090718_Media_alert_Smart_speaker_installed_base_to_hit_100_million_by_end_of_2018.pdf
- Cho, E. (2019). Hey Google, Can I Ask You Something in Private? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (p. 258). ACM.
- Danovitch, J. H., Alzahabi, R. (2013). Children show selective trust in technological informants. *Journal of Cognition and Development*, 14(3), 499–513.

- Davie, N., Hilber, T. (2018). Opportunities and Challenges of Using Amazon Echo in Education. *International Association for Development of the Information Society*.
- Dillon, T. (2018). Creating an Alexa-Enabled Textbook Exercise: An Easy Approach to Custom Automatic Speech Recognition Application. *KOTESOL Proceedings 2018*, 259.
- Dousay, T. A., Hall, C. (2018). Alexa, tell me about using a virtual assistant in the classroom. In *EdMedia+ Innovate Learning* (pp. 1413–1419). Association for the Advancement of Computing in Education (AACE).
- Druga, S., Williams, R., Breazeal, C., Resnick, M. (2017). Hey Google is it OK if I eat you?: Initial explorations in child-agent interaction. In *Proceedings of the 2017 Conference on Interaction Design and Children* (pp. 595–600). ACM. DOI: 10.1145/3078072.3084330
- Gikas, J., Grant, M. M. (2013). Mobile computing devices in higher education: Student perspectives on learning with cellphones, smartphones & social media. *The Internet and Higher Education*, 19, 18–26.
- Griswold, A. (2018). Even Amazon is surprised by how much people love Alexa. (February 2018). <https://qz.com/1197615/even-amazon-is-surprised-by-how-much-people-love-alex/>
- Herrington, A., Herrington, J. (2007). Authentic mobile learning in higher education. In *Proceedings of the AARE 2007 International Educational Research Conference*. Fremantle, Western Australia, 1–9. DOI: 10.1109/ICNICONSMCL.2006.103
- Horn, M. B. (2018). Hey Alexa, Can You Help Kids Learn More?. *Education Next*, 18(2).
- Hoy, M. B. (2018). Alexa, Siri, Cortana, and more: an introduction to voice assistants. *Medical reference services quarterly*, 37(1), 81–88. DOI: 10.1080/02763869.2018.1404391
- Ilhan, K., Mušić, D., Junuz, E., Mirza, S. (2017). Scarlet-Artificial Teaching Assistant. In *2017 International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO)* (pp. 11–14). IEEE.
- Jahan, H. (2017). Reshaping Learning Through Artificial Intelligence. *International Educational Applied Scientific Research Journal*, 2(10).
- Juniper Research. (2017). *Amazon Echo & Google Home to Reside in over 50% of US Households by 2022, as Multi-Assistant Devices Take Off*. <https://www.juniperresearch.com/press/press-releases/amazon-echo-google-home-to-reside>
- Kepuska, V., Bohouta, G. (2018). Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). In *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 99–103). IEEE.
- Kloos, C. D., Alario-Hoyos, C., Muñoz-Merino, P. J., Aguirre, C. C., Castro, N. G. (2019). Principles for the Design of an Educational Voice Assistant for Learning Java. In *International Conference on Sustainable ICT, Education, and Learning* (pp. 99–106). Springer, Cham.4
- Kowalski, J., Jaskulska, A., Skorupska, K., Abramczuk, K., Biele, C., Kopeć, W., Marasek, K. (2019). Older Adults and Voice Interaction: A Pilot Study with Google Home. *arXiv preprint arXiv:1903.07195*.
- Kukulska-Hulme, A. (2019). Intelligent assistants in language learning: friends or foes?. In *World Conference on Mobile and Contextual Learning* (pp. 127–131).
- Lau, J., Zimmerman, B., Schaub, F. (2018a). Alexa, are you listening?: Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW), 102. DOI: 10.1145/3274371
- Lau, J., Zimmerman, B., Schaub, F. (2018b). “Alexa, Stop Recording”: Mismatches between Smart Speaker Privacy Controls and User Needs. Poster at the *14th Symposium on Usable Privacy and Security (SOUPS 2018)*.
- Lei, X., Tu, G. H., Liu, A. X., Li, C. Y., Xie, T. (2017). The insecurity of home digital voice assistants-amazon alexa as a case study. *arXiv preprint arXiv:1712.03327*.
- López, G., Quesada, L., Guerrero, L. A. (2017). Alexa vs. Siri vs. Cortana vs. Google Assistant: a comparison of speech-based natural user interfaces. In *International Conference on Applied Human Factors and Ergonomics* (pp. 241–250). Springer, Cham.
- Lovato, S. B., Piper, A. M., Wartella, E. A. (2019). Hey Google, Do Unicorns Exist?: Conversational Agents as a Path to Answers to Children’s Questions. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children* (pp. 301–313). ACM.
- Malkin, N., Deatrck, J., Tong, A., Wijesekera, P., Egelman, S., Wagner, D. (2019). Privacy Attitudes of Smart Speaker Users. *Proceedings on Privacy Enhancing Technologies*, 2019(4), 250–271.
- McLean, G., Osei-Frimpong, K. (2019). Hey Alexa... examine the variables influencing the use of artificial intelligent in-home voice assistants. *Computers in Human Behavior*, 99, 28–37.
- McReynolds, E., Hubbard, S., Lau, T., Saraf, A., Cakmak, M., Roesner, F. (2017). Toys that listen: A study of parents, children, and internet-connected toys. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 5197–5207). ACM. DOI: 10.1145/3025453.3025735

- Metz, R., (2014). *Voice Recognition for the Internet of Things*. <https://www.technologyreview.com/s/531936/voice-recognition-for-the-internet-of-things/>
- Mulyana, E., Hakimi, R. (2018). Bringing Automation to the Classroom: A ChatOps-Based Approach. In *2018 4th International Conference on Wireless and Telematics (ICWT)* (pp. 1–6). IEEE.
- Neiffer, J. P. (2018). Intelligent personal assistants in the classroom: Impact on student engagement. Ph.D. Dissertation. College of Education and Human Sciences, University of Montana.
- Nilsen, J. H., Røynealand, K. (2019). «It knows how to not understand us!» A study on what the concept robustness entails in design of conversational agents for preschool children (Master's thesis).
- OECD. (2015). *Students, Computers and Learning: Making the Connection*. PISA, OECD Publishing, Paris. DOI: 10.1787/9789264239555-en.
- Pfeifle, A. (2018). Alexa, What Should We Do about Privacy: Protecting Privacy for Users of Voice-Activated Devices. *Wash. L. Rev.*, 93, 421.
- Porayska-Pomsta, K., Alcorn, A. M., Avramides, K., Beale, S., Bernardini, S., Foster, M. E., Frauenberger, C., Good J., Guldborg, K., Keay-Bright, W., Kosyvak, L., Lemon, O., Mademtz, M., Menzies, R., Pain, H., Gnanathusharan, R., Waller, A., Wass, S., Smith, T. J. (2018). Blending human and artificial intelligence to support autistic children's social communication skills. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 25(6), 35. DOI: 10.1145/3271484
- Pradhan, A., Mehta, K., Findlater, L. (2018). Accessibility came by accident: use of voice-controlled intelligent personal assistants by people with disabilities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p. 459). ACM. DOI: 10.1145/3173574.3174033
- Purinton, A., Taft, J. G., Sannon, S., Bazarova, N. N., Taylor, S. H. (2017). Alexa is my new BFF: social roles, user satisfaction, and personification of the amazon echo. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2853–2859). ACM. DOI: 10.1145/3027063.3053246
- Rzepka, C. (2019). Examining the Use of Voice Assistants: A Value-Focused Thinking Approach. *25th Americas Conference on Information Systems*, Cancun, 2019.
- Sayago, S., Neves, B. B., Cowan, B. R. (2019). Voice assistants and older people: some open issues. In *Proceedings of the 1st International Conference on Conversational User Interfaces* (p. 7). ACM.
- Sciuto, A., Saini, A., Forlizzi, J., Hong, J. I. (2018). Hey Alexa, What's Up?: A mixed-methods studies of in-home conversational agent usage. In *Proceedings of the 2018 Designing Interactive Systems Conference* (pp. 857–868). ACM.
- Selak, B. (2017). *Amazon Alexa in the Classroom*. <http://www.coolcatteacher.com/amazon-alexa-classroom/>
- Sheppard, B. (2017). Theological Librarian vs. Machine: Taking on the Amazon Alexa Show (with Some Reflections on the Future of the Profession). *Theological Librarianship*, 10(1), 8–23.
- Skidmore, L., Moore, R. K. (2019). Using Alexa for Flashcard-based Learning. *Proc. Interspeech 2019*, 1846–1850.
- Song, Y. W. (2019). User acceptance of an artificial intelligence (AI) virtual assistant: an extension of the technology acceptance model (Doctoral dissertation).
- Trivedi, N. (2018). ProblemPal: Generating Autonomous Practice Content in Real-Time with Voice Commands and Amazon Alexa. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 80–82). Association for the Advancement of Computing in Education (AACE).
- Tsinonis, T. (2018). How to Use ICT in the Classroom Effectively: The Technological Blend. In *The Future of Innovation and Technology in Education: Policies and Practices for Teaching and Learning Excellence* (pp. 111–125). Emerald Publishing Limited.
- Vtyurina, A., Fournay, A., Morris, M. R., Findlater, L., White, R. W. (2019). VERSE: Bridging Screen Readers and Voice Assistants for Enhanced Eyes-Free Web Search. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 414–426). ACM.
- Wiederhold, B. K. (2018). “Alexa, Are You My Mom?” The Role of Artificial Intelligence in Child Development. *Cyberpsychology, behavior and social networking* 21, 8, 471–472.
- Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th international conference on evaluation and assessment in software engineering* (p. 38). ACM. DOI: 10.1145/2601248.2601268
- Yuan, Y., Thompson, S., Watson, K., Chase, A., Senthilkumar, A., Brush, A. B., Yarosh, S. (2019). Speech interface reformulations and voice assistant personification preferences of children and parents. *International Journal of Child-Computer Interaction*.

G. Terzopoulos is an Adjunct Assistant professor at International Hellenic University and postdoc researcher at the University of Macedonia. His research interests mainly focus on cloud computing, augmented reality and educational mobile applications.

M. Satratzemi is a professor at the Department of Applied Informatics. Her current main research interests lie in the area of Educational Programming Environments and Techniques, Didactics of Informatics, Adaptive and Intelligent Systems, Collaborative Learning Systems, and Serious Games. She has published more than 120 peer-reviewed articles in international journals, conference proceedings. She was Conference co-chair of the 8th ITiCSE (ACM) conference.