

Evaluation of "Speech System" and "Skill": An Interaction Paradigm for Speech Therapy

Vita Santa Barletta^a, Miriana Calvano^b, Antonio Curci^c, Alessandro Pagano^d
and Antonio Piccinno^e

University of Bari Aldo Moro, Bari, Italy

{vita.barletta, miriana.calvano, antonio.curci, alessandro.pagano, antonio.piccinno}@uniba.it

Keywords: e-Health, Smart Assistants, Speech Therapy, User Study, Evaluation, Gamification.

Abstract: Speech therapy is the medical field in which speech impairments are treated. They concern the inability of individuals to adequately enunciate words, construct, elaborate and appropriate sentences when speaking, and overall lack linguistic skills. Although speech impairments can emerge throughout different stages of life, the most common period of time in which they are encountered is childhood. The professionals in this medical field use to treat these impairments with the employment of therapies that are carried out over an extended period of time. This research work aims at proposing and evaluating through a user study a new interactive paradigm that involves "Speech System", a web-application, and "Skill", a skill for Amazon Alexa. The objective consists in determining the practical feasibility of the solution and investigate the consequences that the use of technology brings to the world of speech therapy. The advantages and disadvantages of the interactive paradigm in question are explored and discussed to define the direction of the next steps in this field.

1 INTRODUCTION


Speech therapy is a field of medicine that aims at treating impairments concerning linguistic abilities. De Pompei defines it as "the variety of processes employed by speech-language pathologists who work with the full range of human communication and its disorders. Treatment areas include speech, language, cognitive-communication, or swallowing disorders in individuals of all ages, from infants to the elderly" (DePompei, 2011). Speech impairments can impact individuals in their social, working and academic lives, generating a sense of inadequacy and stress with serious implications. Therefore, experts suggest to make sure to diagnose, recognize and solve them during the early stages of life. The motivation behind this lies in the fact that children are more keen to learning, acquiring new skills, and correcting behaviors. This case study considers speech therapy as a field that involve three actors: speech therapists, caregivers, and patients.


Therapists. They make diagnoses and create and manage therapies that are administered to patients, who are monitored and assessed during the whole process to understand improvement levels and/or change the direction of the treatment, if necessary. Therapies consist of personalized exercises whose difficulty is weighted on the level of severity of the patient's disorder.


Caregivers. Caregivers play a crucial role in speech therapy because they are responsible of guiding and assisting children in attending appointments and following through with the treatment, even outside medical facilities. They help patients when they have to practice at home and provide emotional support, acting as a middleman (Barletta et al., 2023b).


Patients. Patients are the subjects of speech therapy and, in this research work, they are children whose age ranges from 4 to 8 years old. They are assigned exercises and need to carry them out to improve their condition and eventually solve their impairments.


The introduction of technology in speech therapy has the goal of supporting the actors involved in the process by reducing the cognitive demand of tasks

^a  <https://orcid.org/0000-0002-0163-6786>

^b  <https://orcid.org/0000-0002-9507-9940>

^c  <https://orcid.org/0000-0001-6863-872X>

^d  <https://orcid.org/0000-0002-7465-9778>

^e  <https://orcid.org/0000-0003-1561-7073>

that can be automatized and making them feel more comfortable.

Using Artificial Intelligence (AI), employing Smart and Internet of Things (IoT) devices, undertaking a Human-Centered Design (HCD) approach, and applying gamification can make the difference in the way that all the parties involved in speech therapy interact with each other and perform their activities. More specifically, AI can be an incredibly powerful tool when it comes to automatize tasks that require unnecessarily high costs in terms of resources and time, such as the correction of exercises or the diagnosis of patients based on their performance. According to Dorsemayne et al., IoT is a "group of infrastructures interconnecting connected objects and allowing their management, data mining and the access to the data they generate", which implies that different objects pertaining to the same environment can cooperate and collaborate to facilitate daily tasks and support individuals in their activities, enabling remote controlling, vocal interactions, customization, and so much more (Dorsemayne et al., 2015). On the other hand, the treatment process can be a stressful and frustrating process for children, because they are constantly required to step outside of their comfort zone and be exposed to potential failure. Therefore, gamification, which is "the use of game design elements in non-game contexts" (Groh, 201), can change children's perception and make them feel more at ease, enjoying the tasks that they are assigned, enhancing engagement levels. Therapies can become a more pleasurable and playful experience, making it possible to achieve learning goals in shorter time spans and with higher accuracy (Desolda et al., 2021).

In this research work, an interaction paradigm is proposed, which combines a web application, called "*Speech System*", and a skill designed for Alexa¹, the Amazon's Smart Assistant, called "*Skill*". The aim of the proposed solution is to allow therapists to manage treatments and patients in a continuous, hybrid, and systematic way. More specifically, it has the goal of enabling patients to perform exercises at home avoiding to waste resources in travelling to attending physical appointments at the doctor's office, while benefiting from the advantages and guidance that can be earned from the employment of smart objects and voice assistant (Barletta et al., 2023a; Calvano et al., 2023). The interaction paradigm will be later explored and illustrated.

The HCD approach states that users have to be involved in every phase of the creation of a new product, service or technology; testing and appropriately evaluating the outputs of any product can allow to

reach higher satisfaction levels of users when interacting with it, while guaranteeing more effectiveness and efficiency (NIST, 2021). User studies are crucial to assess the feasibility and validity of the work performed and to determine its positive and negative aspects. Therefore, a user study was conducted, providing useful insights into the interaction paradigm and how the users perceived it.

2 RELATED WORK

Speech therapy can be integrated into smart homes thanks to the numerous tools and technologies available today (Barletta et al., 2022). Smart voice assistants, such as Amazon Alexa, Google Assistant, and Apple HomePod, can be exploited in the medical field to introduce more consistent and daily support. The reason lies in the fact that treatments and therapies need consistency, which is an aspect that can be reached with daily practice and the right medical support. It is important for professionals to be able to follow their patients throughout the entire process: from diagnosis to disorder solution or improvement; it becomes necessary to have remote monitoring tools at their disposal rather than only traditional in-person appointments.

In this context, Qiu et al. explores how voice assistants can enable remote delivery of speech therapies at scale; in particular the objective of this study consists in helping individuals with speech impairments to have access to treatment in case of absence of trained professionals and available resources (Qiu and Abdullah, 2021).

In addition, Cassano et al. and Buono et al. propose scenarios in which smart home devices can be used to support speech therapy for children with linguistic impairments and how different interconnected devices, with the help of End-User Development techniques and thanks to higher levels of engagement, can increase the chance of success of treatments in this field. Therefore, it states that IoT devices can be used to administer therapies and support therapists (Buono et al., 2019; Cassano et al., 2019). Moreover, it has been determined that the employment of a vocal assistant in smart home environments can present interesting challenges, such as the presence of environmental noise and how the integration of Automatic Speech Recognition systems can be helpful in the scenario (Aman et al., 2016).

Another important aspect to highlight is that automation of the home environment is not the only factor that improves and facilitates the therapy process, but also speech recognition is crucial. In this regard,

¹<https://alexa.amazon.com/>

Chern et al. propose a smartphone-based hearing assistive system that includes voice-to-text conversion to make speech recognition easier, as smart voice assistants can be used to provide personalized feedback that adapts the system to the unique needs of the user (Chern et al., 2017).

An additional study on this issue was conducted by Lecouteux et al. and Parameshchhari et al., where other Automatic Speech Recognition techniques are explored in more detail about the artificial intelligence algorithms behind the mentioned technology; it has the potential to support speech therapy and improve communication for individuals with speech disorders (Lecouteux et al., 2011; Parameshchhari et al., 2013).

Furthermore, the use of speech recognition models in speech therapies has also been studied by introducing a correlation with the concept of serious games. First of all, it is important to explain the concept of serious games, which can be defined as 'games that do not have entertainment, enjoyment, or fun as their main purpose' (George, 2019).

For example, Ganzeboom et al. and Vogel et al. explore the use of speech recognition technology in speech therapy; they find that the introduction of serious games, by triggering intrinsic motivation and participation of patients, can improve the effectiveness of speech therapies, especially in patients with dysarthria due to Parkinson's disease (Ganzeboom et al., 2022; Vogel et al., 2022).

On the other hand, it is common to think that the typical target users of these models are younger people, but this is wrong. In fact, another important contribution has been made by analyzing these models and also considering the perception of older people (Aman et al., 2016; Werner et al., 2023). In fact, the main contribution consists of highlighting the differences between people who are users of the system that involve speech recognition and those who are not; the authors determined that the non-users see their age as an obstacle to this type of technology and showed inhibition to try it. Otherwise, users find it easy to use, but have raised concerns about the transparency and privacy of their personal data (Werner et al., 2023).

In addition, systems and applications with the goal of facilitating language development can be found in the literature. Some examples are indicated below:

- **Happi Scrive**² and **KidEWords**³ are crossword puzzle applications for children that help them develop writing skills.

²<https://apps.apple.com/it/app/happi-scrive/id464675842>

³<https://apps.apple.com/it/app/kidewords-by-chocolaps/id879490139>

- **Teach and Touch**⁴ is an application through which speech therapists offer children the opportunity to perform a personalized rehabilitation path on specific morphosyntactic difficulties.
- **Training Cognitivo**⁵ is a project with the objective of evaluating and training language disorders that affect preschool and school children, adolescents, and young adults. In this initiative, a group of speech therapists and psychologists is involved.

Finally, it emerges that the potential advantages that can be gained from exploiting the mentioned technologies are worth the effort of designing and developing a web application that possesses gamification and game-based learning elements to keep track of therapies, follow children in their treatment and allow them to carry out exercises from the comfort of their home.

3 INTERACTION PARADIGM

This study involves three actors (i.e. speech therapists, caregivers and patients) and a home equipped with smart devices. In the interaction paradigm, the process takes place in a smart home, in which the patient and the caregiver live, through the employment of smart assistants which allow users to interact vocally. The latter, represents an external entity that enters the scene to solve problems and guiding the user during the process. On one side there is the therapist who interacts with the web application and uses its functionalities to perform therapy management activities and, on the other side, there are the patient and the caregiver who interacts both with the smart assistant and the web application (Barletta et al., 2023a). It is important to note that the actors involved in this study are the same as those considered in the work previously cited.

3.1 "Speech System": A Web-Application

"Speech System" is a web application that provides three separated personal areas belonging to each actor involved in the process; Figure 1 illustrates the welcome page of the application.

In the therapist section, the system provides multiple functionalities to allow professionals to create diagnoses, use or create exercise, and employ them to create therapies to administer to patients. Three types of exercises are featured:

⁴<http://www.teachandtouch.it/>

⁵<https://www.trainingcognitivo.it/>

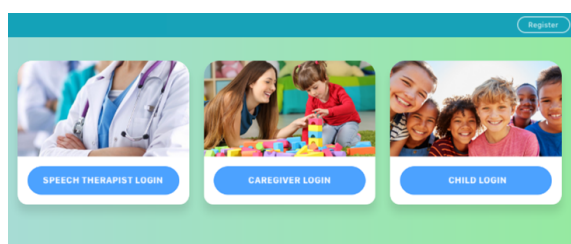


Figure 1: "Speech System" and its three sections for each actor.

- **Naming Images:** the patient has to identify the objects represented in the images by enunciating their name.
- **Repetition of Words:** A set of words characterized by similar phonetic characteristics is shown to the patient who have to enunciate them.
- **Minimum Pair Recognition:** The patient has to click on the loudspeaker button and, after, "Speech System" will vocally indicate the object on which he/she has to click.

Gamification is embedded in the child's section of the system: after each exercise the patient gains cookies to reinforce the reward mechanism and increase engagement. The exercises performed are recorded and available for the monitoring of the speech therapist and the caregiver.

The caregiver's section allows them to view appointments, start or stop therapies that were assigned to their child, and customize the graphical interface to make their personal area more welcoming and adjusted according to their preferences.

3.2 "Skill": An Amazon Alexa Skill

As mentioned previously, the interaction paradigm encompasses "Skill", a skill for Amazon's Alexa developed ad-hoc for the purpose of this study.

Alexa is a smart assistant which is characterized by machine learning algorithms throughout it is possible to develop vocal performances, which offers to the user a more intuitive way to interact with technological devices. More specifically, using the Amazon's Skill Kit and Developer Console⁶, developers can create new functionalities, called 'Skills'. In addition, each skill is characterized by an intent, which are actions that the device has to perform to satisfy the user's requests and are launched through sample utterances defined by the developer while implementing the skill. Sample utterances contain input parameters, called slots, with the objective to personalize the

⁶<https://developer.amazon.com/alexa/console/ask>

action performed by the smart assistant according to the user's needs. The skill in question is characterized by three main functionalities, which allow the user to launch the skill, to create a reminders for the patient to perform the exercises of the day, and, finally, to start the therapy. A detailed explanation of the features in question is provided below.

Skill Launch. To use this feature, the user has to enunciate the wake-up phrase "Alexa launch "Skill"". Then, the smart assistant answer with the expression "Hello <NameOfTheChild>! Tell me "I am at home" to start therapy" to welcome they.

Reminder Creation and Usage. Reminders are used by caregivers to define when and where the child must perform the therapy. With the objective to make the interaction with Alexa more flexible, multiple sentences that the user can pronounce to create the reminder were created. In particular:

- Create a reminder on [date] at [time] in [room].
- Set a reminder [date] at [time] in [room].
- Create a reminder [date] at [time] in [room].
- Set a reminder on [date] at [time] in [room].

It is important to highlight that the words "time", "date", and "place" are slots that saved in the database and are the field that allow the user to personalize the reminder. Consequently, if the reminder is set successfully, the smart assistant gives the feedback to the user saying "The reminder was successfully set. " and at the time and place for which the reminder was created, the speaker automatically activates and says "It is time to start your therapy! Launch the skill and tell me "I am at home"".

Therapy Initiation. The patient is able to start to perform the therapy activities by pronouncing the following expression "I am at home". After, the smart assistant will indicate to the user to go in the room setting while creating the reminder saying, for example, "Go to <RoomSetByCaregiver>! If you are already there, tell me 'I am here'!". Now, the smart assistant, after giving the instruction that are necessary to find the place in which the tratment will be performed, is waiting for the answer from the user. Finally, the smart assistant is ready to let the child start the therapy and, thus, says "Let's start!". In order for this feature to work properly, there must be a therapy available in e-SpeechT, administered by the therapist and started by the caregiver.

3.3 A Stereotypical Scenario

In this section, how the integration of "Speech System" with Amazon Alexa works is presented. As shown in Figure 2, the scenario in question that takes into account a stereotypical structure of a common house, (Barletta et al., 2023b).



Figure 2: Example of Smart Home (1. Caregiver's Bedroom, 2. Kitchen, 3. Living Room, 4. Child's Bedroom, 5. Bathroom).

From the Figure 2, it is possible to notice that smart assistants are set in the living room, kitchen, and child's bedroom. In this context, it is important to underline that the role of the caregiver is crucial because they have the responsibility of setting up the environment to prepare the child to perform the activity required by the therapy.

The interaction with the smart assistant works in the following way:

1. The caregiver sets a reminder through Alexa for a specific date and time.
2. At the indicated moment, Alexa wakes up and says "It's time to start the therapy! Go to the <RoomSetByCaregiver>. If you are already there, tell me 'I am here!'".
3. The child follows the Alexa's indications and replies to the smart assistants with the suggested sentence. After this, the patient goes to the indicated room.
4. Alexa enunciates the following expression "Let's start!" and launches "Speech System" on the patient's device.

In conclusion, it is underlined that this interaction paradigm helps the patients to be more autonomous while performing the exercises assigned by the speech therapist, even in absence of the caregiver. The employment of the smart assistant represents also a way

to introduce gamification and gaming elements in the process to increase their engagement and motivation levels.

4 USER STUDY

This section explores the planning and execution phases of the user-usability study conducted to test the interaction paradigm in question and how patients behave while performing the therapy at distance. The results are discussed and commented, highlighting the problems, weaknesses, and strengths points.

4.1 Planning

This user study was planned to be conducted in the presence of a facilitator and two observers, who are responsible for making the participants feel welcomed and explaining how the study is carried out. In addition, each participant was required to carry out the tasks following the "thinking aloud" technique; it consists of asking to talk and express aloud how he/she is feeling, what difficulties are being encountered, the doubts that are emerging, and whether the goal of the task is thought to be reached or not. Verbalization by users makes it possible to understand not only what problems the user experiences in using the system, but also why said problems arise. It is, moreover, an excellent system for obtaining a large amount of information with the participation of a small number of subjects.

The objective of the study was to test the usability and accessibility of the system and its integration with Amazon Alexa voice assistance; in fact, the goal was not to evaluate the medical effectiveness of the administration of therapies through "Speech System", but rather to see how children interact with the smart assistant with respect to ease of use when performing exercises in the application. The target users of this study were children aged 4 to 8 years, which gave the possibility of testing the services provided by the application for users with different linguistic capabilities and cognitive maturity, keeping in mind the target audience for which the system was originally designed. Being a pilot study for this interaction paradigm, the suitable number of participants was considered between eight and ten for reliable results; therefore, the actual execution of the study involved 10 children, encompassing all the characteristics of the target users in question. The participants were chosen through convenient sampling (Bellhouse, 2005). Each test was carried out in an informal setting, with the aim of ensuring as little stress as possible for the participants,

as children can distract themselves or get upset quite easily when under pressure or in uncomfortable situations.

The actual link between Amazon Alexa and "Speech System" was performed with the Wizard of Oz technique, which consists of realizing the prototype in which the participants interact with the system where answers and feedback are provided by the experts behind the scenes, unbeknownst to them. The feature to be simulated through this technique was the opening of the application right after the final phrase played through the speaker; more specifically, one of the experts opens the browser on a computer before the children approach the device.

A crucial part of studies with users are questionnaires. To administer them correctly, it is necessary to measure the questions based on the individuals to whom they are addressed. Your age is relevant since it determines the attention curve, your memory skills, and your ability to adapt your response to the context.

In light of the latter, two types of questionnaire were designed for the study: First, a Google Form⁷ questionnaire which has the objective of evaluating the consistency of the state of concentration of the child, the ability of Alexa to recognize vocal commands from the children, and the enjoyment of the interaction with both "Speech System" and Amazon Alexa. It is administered to the observer of the test, who acts as a caregiver. It contains questions with a Likert scale, multiple choice, and open answers, depending on the topic and the subject.

The choice of questions with different responses was made to keep the individuals engaged and prevent them from responding in a distracted and unfocused way. As shown below, the questionnaire is divided into two sections:

- The first seven questions are directed at the caregiver's experience while using the application.
- the other addresses specific issues regarding the patient.

Caregiver:

1. Were problems encountered when starting the therapy?
2. Indicate the specific problems in case you answered "Yes" to the previous one.
3. Do you think that the child would have completed all of the exercises in your absence?
4. When analyzing the statistics of exercises, do you think that the automatic correction works accurately and correctly?

⁷<https://www.google.it/intl/it/forms/about/>

5. Was it hard for you to follow the child during the exercise execution?
6. On which device was the system used?
7. Did you encounter other specific difficulties? If so, please list them.

Child:

1. How long did the child use "Speech System"?
2. Did the child ask for help?
3. Did the child's mood negatively change during the execution of the exercises?
4. In case you replied "Yes" to the previous question, why do you think the mood changed?
5. The child showed signs of anger and irritability while using "Speech System".
6. The child remained concentrated for the entire duration of the game.
7. Did the child want to continue playing?
8. The child perceives the game as: an obligatory task, a game
9. Which difficulties were encountered when logging onto the system?
10. The child got bored when using the system.
11. The child wanted to stop the execution of the exercises in advance.

This questionnaire aims at collecting both qualitative and quantitative results, such as the user's difficulties during the task execution and the percentage of task success, respectively. For example, questions 1) and 2) are relevant to understand whether the login phase was problematic for the child or not. Questions 3) and 9) aims at measuring the effectiveness of the application and whether the child is autonomous. Questions 10-12, 17) were administered to get an insight into the mood of the child to establish the positive or the negative implications coming from the integration of the proposed technology. This aspect is crucial for the study because anger and stress can cause the medical ineffectiveness of the exercises.

In conclusion, to understand whether the goal of masking therapy as a playful experience was met or not, Question 15) is administered. The second is a single-question questionnaire, which has the goal of directly asking the child how he/she feels after the test. The answer is provided through a smileyometer (Bell, 2007; Jesus et al., 2019), as shown in Figure 3 in order to let participants of all ages be able to reply; the question is "Congratulations! You completed your exercises. How do you feel?".

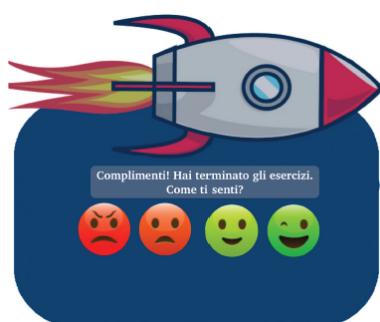


Figure 3: Single-question questionnaire for children through a smileyometer.

Moreover, it also planned a small reward for each child who takes part in the study. The necessary material and pieces of technology needed to successfully perform the study are a computer to use "Speech System", a smart speaker, and a tablet for the smileyometer question. The motivation behind this choice lies in the fact that the Alexa Skill to be tested is still under development and has not been released to the public.

4.2 Execution

After the completion of the planning phase, participants were recruited with respect to the requirements previously established. Considering that the target audience of the system was children aged 4 to 8 years, the permission of the parents and caregivers was necessary to proceed; their presence was crucial for the execution of the tests because they were conducted on-site, at the participants' homes, with their supervision.

The study involved 10 children, 7 girls, and 3 boys. It was crucial for the children to be comfortable with the environment in which they were tested, to reduce any chance of additional stress or frustration during the study, as seen in the Figure 4.

The experts made it sure to set "Speech System" beforehand, by creating the profiles for each child, new therapies, and exercises, which were necessary to test the system simulating a real-life scenario. The facilitator and study observers set all devices in the environment while each child was distracted, making the experience more immersive for them.

More specifically, each caregiver was asked in which room the child would have had to perform the test, in order to program the Amazon Alexa Skill accordingly. In the latter room, a computer was set with the starting page of "Speech System", ready to be used by the children, as if the voice assistant opened it. Before starting with the actual test, the context and goal of the study were carefully explained to each child, taking into account their age and making it all seem



Figure 4: Picture of one of the children using "Speech System".

as playful and light as possible. Each child was then accompanied to a room in their home where the Amazon Echo was set, and then the test began.

During each test, observers paid attention and took notes of the child's movements, emotions, and what he was saying. One of the experts acted as a caregiver and was there in case the child needed help or guidance throughout the process.

In order to understand the effectiveness of the invocation phrase of the skill, two options were the subject of study: "Launch Skill" and "Launch Therapy"; they were tested through the within-subject design. When the child completed all the exercises in "Speech System", the smileyometer question was administered. Meanwhile, one of the observers filled in the Google Form questionnaire, acting as their caregiver.

4.3 Results and Discussion

In this section, the results obtained during the study will be presented. More specifically, the analysis of participants' behavior during the tests and their responses to the questionnaires will be explored, drawing conclusions from them.

During the tests it stood out that this interaction paradigm involving Alexa and "Speech System" cannot be a one-size-fits-all solution.

The skills and capabilities of children develop at a very fast pace, which inevitably creates a gap between those of 4-5 and 6-8 years old, caused by the reading and writing skills or the lack of them. Nevertheless, the experts, being already aware of this issue, were ready to support the children by reading words and phrases on the screen when necessary. At the same time, it was noticed that the presence of an adult figure that gave hints about what to do made the participants feel under pressure and examination, leading to a state

of inadequacy and discomfort.

On the other hand, older children were already familiar with smart assistants and computers due to their employment at home and in schools, which made it easier to perform the test and explain to them how to carry out the tasks; it led to more autonomy during all phases of the study.

Even though every child wished for a different invocation phrase for the skill in question, a further distinction can be made: the younger children could not even enunciate the initial one, as opposed to the older ones. It becomes clear that it is necessary to create ramifications of "Speech System" and its integration with Alexa, depending on the characteristics of its user. More specifically, two aspects to focus on in order to improve this interaction paradigm are: choosing a different invocation phrase that recalls the concept of playing as mentioned in the previous paragraph and suggested by the children; differentiating the types of exercises.

Nevertheless, it is important to underline that a factor kept in mind throughout the whole process of analyzing the results was the embarrassment and shyness of the participants, which can cause reticence. This aspect is considered as a potential bias. In conclusion, it can be asserted that in spite of the previously mentioned problems, the overall experience was perceived as enjoyable and fun and something to repeat in the future. This suggests that technology can be a great tool for helping children in therapy by transforming boring and serious activities into something fun and playful.

5 CONCLUSIONS AND FUTURE WORK

The employment of technology in medicine has been spreading at a very fast pace, expanding the horizons for every party involved. When it comes to speech therapy, e-health enables professionals to assist patients remotely, monitor in a more systematic way the therapies that they administer, and automatize some of their tasks. At the same time, children can engage in less stressful activities with the use of gamification techniques and game-based learning. In addition, by not attending recurrent physical appointments, patients are less exposed to feelings of frustration and anxiety due to situations that can be perceived as exams or tests.

In this research work, these concepts are embodied in "Speech System" and "Skill", converging towards a new interaction paradigm that connects and intertwines the advantages of a web-application and

the employment of smart assistants.

The analysis of the results made it possible to ascertain that, broadly speaking, children enjoy the system and the interaction with Alexa to the extent that they wished to continue playing. However, even if the activities were perceived as funny and pleasant in the majority of cases, some problems were encountered that deserve to be studied in-depth, such as the difficulty for children to enunciate the invocation phrase, following the instructions to perform some exercises in "Speech System" that required more advanced reading or speaking skills. Future work is intended to involve a more specific catering of the exercises in "Speech System" to children depending on their age; older children performed activities too easily, as opposed to the younger ones who were unable to read and with lower linguistic skills.

In addition, it is planned to include the employment of a metric to further standardize the study when evaluating it.

Lastly, a future work that is intended to perform concerns the execution of a longitudinal study to assess the impact of the interaction paradigm from a medical point of view, too. This activity can provide useful insights on the strengths and weaknesses of the proposed solution.

ACKNOWLEDGEMENTS

The work was partially supported by the SMATH4SD project (Smart Therapies for Speech Disorders) under the grant 'Horizon Europe Seeds 2021' funded by the University of Bari.

REFERENCES

- Aman, F., Aubergé, V., and Vacher, M. (2016). Influence of expressive speech on asr performances: Application to elderly assistance in smart home. In Sojka, P., Horák, A., Kopeček, I., and Pala, K., editors, *Text, Speech, and Dialogue*, pages 522–530, Cham. Springer International Publishing.
- Barletta, V., Calvano, M., Curci, A., and Piccinno, A. (2023a). A Protocol to Assess Usability and Feasibility of e-SpeechT, a Web-based System Supporting Speech Therapies. In *Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies*, pages 546–553, Lisbon, Portugal. SCITEPRESS - Science and Technology Publications.
- Barletta, V. S., Calvano, M., Curci, A., Piccinno, A., and Pagano, A. (2023b). Poster: Speech Therapies and Smart Assistants: an interaction paradigm proposal. In *Proceedings of the 15th Biannual Conference of*

- the Italian SIGCHI Chapter*, pages 1–3, Torino Italy. ACM.
- Barletta, V. S., Cassano, F., Pagano, A., and Piccinno, A. (2022). New perspectives for cyber security in software development: when end-user development meets artificial intelligence. In *2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, pages 531–534. IEEE.
- Bell, A. (2007). Designing and testing questionnaires for children. *Journal of research in nursing*, page 462.
- Bellhouse, D. R. (2005). *Systematic Sampling Methods*. John Wiley & Sons, Ltd.
- Buono, P., Cassano, F., Piccinno, A., and Costabile, M. F. (2019). Smart objects for speech therapies at home. In Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M., and Zaphiris, P., editors, *Human-Computer Interaction – INTERACT 2019*, pages 672–675, Cham. Springer International Publishing.
- Calvano, M., Curci, A., Pagano, A., and Piccinno, A. (2023). Speech therapy supported by ai and smart assistants. In *International Conference on Product-Focused Software Process Improvement*, pages 97–104. Springer.
- Cassano, F., Piccinno, A., and Regina, P. (2019). End-user development in speech therapies: A scenario in the smart home domain. In Malizia, A., Valtolina, S., Morch, A., Serrano, A., and Stratton, A., editors, *End-User Development*, pages 158–165, Cham. Springer International Publishing.
- Chern, A., Lai, Y.-H., Chang, Y.-P., Tsao, Y., Chang, R. Y., and Chang, H.-W. (2017). A smartphone-based multi-functional hearing assistive system to facilitate speech recognition in the classroom. *IEEE Access*, 5:10339–10351.
- DePompei, R. (2011). *Speech-Language Therapy*, pages 2343–2344. Springer New York, New York, NY.
- Desolda, G., Lanzilotti, R., Piccinno, A., and Rossano, V. (2021). A System to Support Children in Speech Therapies at Home. In *CHIItaly 2021: 14th Biannual Conference of the Italian SIGCHI Chapter*, pages 1–5, Bolzano Italy. ACM.
- Dorsemaine, B., Gaulier, J.-P., Wary, J.-P., Kheir, N., and Urien, P. (2015). Internet of Things: A Definition & Taxonomy. In *2015 9th International Conference on Next Generation Mobile Applications, Services and Technologies*, pages 72–77, Cambridge, United Kingdom. IEEE.
- Ganzeboom, M., Bakker, M., Beijer, L., Strik, H., and Rietveld, T. (2022). A serious game for speech training in dysarthric speakers with parkinson’s disease: Exploring therapeutic efficacy and patient satisfaction. *International Journal of Language & Communication Disorders*, 57(4):808–821.
- George, K. (2019). *Educational game design fundamentals : a journey to creating intrinsically motivating learning experiences / George Kalmpourtzis*. Taylor & Francis Group.
- Groh, F. (201). Gamification: State of the art definition and utilization. *Research Trends in Media Informatics*, pages 39–46.
- Jesus, L., Santos, J., and Martinez, J. (2019). The table to tablet (t2t) speech and language therapy software development roadmap. *JMIR Research Protocols*, 8.
- Lecouteux, B., Vacher, M., and Portet, F. (2011). Distant speech recognition in a smart home: Comparison of several multisource asrs in realistic conditions. In *Interspeech*.
- NIST (2021).
- Parameshachari, B., Gopy, S. K., Hurry, G., and Gopaul, T. T. (2013). A study on smart home control system through speech. *International Journal of Computer Applications*, 69(19).
- Qiu, L. and Abdullah, S. (2021). Voice assistants for speech therapy. *UbiComp ’21*, page 211–214, New York, NY, USA. Association for Computing Machinery.
- Vogel, A. P., Graf, L. H., Magee, M., Schöls, L., Rommel, N., and Synofzik, M. (2022). Home-based biofeedback speech treatment improves dysarthria in repeat-expansion scas. *Annals of Clinical and Translational Neurology*, 9(8):1310–1315.
- Werner, L., Huang, G., and Pitts, B. J. (2023). Smart speech systems: A focus group study on older adult user and non-user perceptions of speech interfaces. *International Journal of Human-Computer Interaction*, 39(5):1149–1161.