A Protocol to Assess Usability and Feasibility of e-SpeechT, a Web-based System Supporting Speech Therapies

Vita Santa Barletta\textsuperscript{a}, Miriana Calvano\textsuperscript{b}, Antonio Curci\textsuperscript{c} and Antonio Piccinno\textsuperscript{d}

\textit{Computer Science Department, University of Bari Aldo Moro, 70125 Bari, Italy}

Keywords: Speech Therapy, e-Health, Speech Impairment, Gamification, Usability Study, Telemedicine, Game-based Learning.

Abstract: Speech disorders and impairments are a heterogeneous group of dysfunctions that concern linguistic abilities and, in the majority of cases, they affect people during their childhood from 4 to 6 years old. Appropriate treatment should be defined for each patient according to their problems, which can be physical, such as muscle weakness, brain damage, vocal cord damage, or paralysis, and psychological, such as autism, PTSD or Down Syndrome. Therefore, considering these aspects, this research work aims at supporting patients in carrying out therapies to solve speech impairments through gamification, game-based learning, e-health, and telemedicine. A web-based system is proposed to achieve the goals by supporting all stakeholders involved in speech therapy. The design and plan of a usability and a longitudinal study have been presented, with the goal of testing the system’s usability and medical efficacy by involving real end-users and domain experts.

1 INTRODUCTION

Speech therapy is the treatment of speech disorders founded on the application of notions and methods pertaining primarily to phoniatrics but also to other disciplines such as glottology, psychology, pedagogy, and neurology (Law et al., 2003).

Speech disorders refer to the disruption of “normal speech”: they affect the vocal cords, muscles, nerves, and other structures within the throat, due for example to vocal cord damage, brain damage, muscle weakness, or vocal cord paralysis. So, the areas of communication involved in this disorder are phonology, the pattern of sounds used by the child; grammar, the way that language is constructed; vocabulary, the words that a child can say and understand; morphology, meaningful changes to words to signal tense, number, etc.; narrative skills, the ability to relate a sequence of ideas; pragmatic language that refers to the social language skills in the interaction with others and in a given situation (Hourcade, 2015; Law et al., 2003; Adams et al., 2012).

Nowadays, rehabilitation therapies for speech problems are administered by specialists, who are called speech therapists, and consist in the performance of exercises calibrated and chosen with the aim of solving the specific problems of each subject. Assessment of eligibility for intervention includes a combination of standardized assessment, observations of linguistic and communicative performance, and professional judgment (Law et al., 2003). Thanks to speech therapies, speech disorders can improve: in order to strengthen the muscles of the face and throat, a professional therapist guides the patient through different types of exercises (Barletta et al., 2022). It is important to underline that therapists prescribe children to also practice also at home. In some cases, exercises are perceived as boring, which causes children to underperform, to carry them out incorrectly, or to refuse to perform them. Unfortunately, this aspect is not verifiable by therapists and it is out of their control. Thanks to the currently available technology, these limitations can be overcome and patients can feel more comfortable.

In speech therapy, three categories of stakeholders are involved:

1. Speech therapists, who are doctors who have the goal of resolving problems involving speech impairment and disorders. They usually define a personalized treatment program to help solve children’s speech difficulties and prevent subsequent
literacy problems.

2. Caregivers, who are parents or anyone who looks after the children when they have to follow speech therapies.

3. Patients, who are children from 4 to 6 years old who are affected by speech disorders.

Currently, limited technological applications support speech therapists and patients during therapy at home (Cassano et al., 2022). But at the same time, the researchers demonstrate that the idea of supporting this process by using multimedia and smartphone to make the activity more enjoyable for the target users is quite widespread (Jesus et al., 2019a; Pentiuc et al., 2010; Rossano et al., 2020). Thus, considering the ever-evolving and ever-growing ICT in the medical field, the idea is to develop a system that simplifies the process of administering therapies to children, because it enables them to carry out exercises from their homes, having their caregiver’s supervision and the therapists’ remote monitoring (Desolda et al., 2021; Butcher et al., 1987). This can be possible if therapies are delivered through mobile and “smart” computing devices. This results in the development of mobile applications or, even better, the introduction of systems that can be accessed from mobile devices, and with the introduction of new concepts and principles, such as “gamification” and “game-based learning” (Saeedi et al., 2022) for a more effective employment in following the therapies.

The paper presents a web-based system to support speech therapists in administering and monitoring therapies that aim at solving problems involving phonology, morphology, content, and structure of speech. It represents a concrete application of principles, such as e-health, telemedicine, gamification, and game-based learning. The system targets are children as patients, from 4 to 6 years old, speech therapists, and caregivers.

The goals are to help therapists to manage and to monitor their patients in a quicker and more structured way; patients carry out the exercises without the need to physically go to a facility. In addition, the plan of a usability study and a longitudinal study is presented, which will be performed in the next future to test the system’s usability and medical efficacy.

2 SYSTEM DESCRIPTION

e-SpeechT is a system that offers different functionalities and graphical interfaces for each type of user will be shown. On the welcome page of e-SpeechT, users can select their role, log in or create a new account. Therapists, caregivers, and patients play three distinctive roles, according to which the system displays and enables different functionalities. In order to be consistent with its diverse audience, when children use the application, the graphical interface becomes more playful.

2.1 Speech Therapist

The speech therapist’s perspective of e-SpeechT allows to create and to monitor therapies for each child.

More specifically, as shown in Figure 1, the main actions that the system allows speech therapists to perform are listed below: view current and previous diagnoses; create new therapies and view previous ones; view the results and their statistics of the exercises carried out by the patient; check and perform corrections on the previously-mentioned exercises; view patients’ personal data.

Figure 1: An example of the patient’s web-page from the therapists point of view.

The menu section in the upper-central part of Figure 1 is composed of four sections:

1. add a new patient;
2. monitor and create exercises for therapies;
3. open a calendar where the schedule of appointments for each patient is shown;
4. receive miscellaneous notifications.

In addition, speech therapists can create a series, that is a predefined sequence of exercises that will be administered each time in the same order. It is crucial to adapt the level of difficulty of exercises to the patient’s needs, age, and diagnosis in order to make a balanced and proper evaluation (Pagano and Marengo, 2021). e-SpeechT allows therapists to set an error tolerance threshold that plays an important role in the automatic correction of exercises performed by the system. This functionality was implemented by training a classification model using the voice recordings of fifteen children. In Figure 2, an example of threshold setting is shown.
2.2 Caregiver

Similarly to speech therapists, the first page shown to caregivers is an overview of all the children that they are associated with, because there might be multiple patients in the same family. The following functionalities are offered to the caregiver: accessing the calendar section where all the appointments and medical examinations available; accessing the gifts and rewards section; viewing reports and statistics of the exercises performed by the child; viewing the child’s personal information; changing the aspect of the child’s personal area by setting a playful wallpaper, called “scenario”.

Through their caregiver’s account, children can set the password to log into their personal area. Since this system targets children aged from 4 to 6 years old, their reading and writing skills are at a very early stage or not present at all, therefore, a graphical password has been opted for. In order to let children feel comfortable and welcomed, the graphical password consists of a cartoon character to be chosen from a list proposed by the system.

Caregivers are allowed to start or stop the administration of therapy in order to choose a specific and dedicated period of time in which a child has to carry out the designated exercises. This functionality is shown in Figure 4.

On the top right, a drop-down menu can be opened in which caregivers can view their personal information and log out of the system.

2.3 Patient

Patients use the system to carry out exercises prescribed by the therapist by playing serious games. Examples of exercises are shown, described, and analyzed in the next section.

3 THE GAME-BASED LEARNING APPROACH

The e-SpeechT system embodies principles like gamification and game-based learning. Every aspect of the system that involves interaction with children is in the form of games or contains playful aspects (Di Bitonto et al., 2014). In the following sections, the patient area will be illustrated.

3.1 Login-Method

The login phase is composed of two steps: speech therapist choice and password selection; each patient can access their personal area by choosing among the
list of speech therapists in the system by recognizing their profile picture; then, they will need to select the password image chosen together with her/his caregiver, as shown in Figure 5.

Figure 5: Child Login page: therapist and password selection.

Once the child has successfully logged in to the system, a home screen is shown and both a textual and audio welcome message is given.

3.2 Exercises

Three categories of exercises can be identified in e-SpeechT:

1. **Naming images**;
2. **Minimum pair recognition**
3. **Repetition of Words**.

For the first type of exercise, Naming Images, the system will show on the left the image associated with the word to pronounce and on the right an orange button with the icon of a microphone. Under this button, hints are provided for the execution of the exercise in the form of buttons representing a loudspeaker; press them, the system will reproduce vocally the selected hint. See, as an example, Figure 6.

In the Minimum Pair Recognition exercise, the graphic interface will consist of two images corresponding to the words placed on the sides of the screen, and a button with a speaker icon between them: by pressing this button, the system will reproduce the name of the correct word. The position of the correct image will be randomly generated each time a user accesses the execution of the exercise. To perform the exercise correctly, the patient must reproduce the name of the word, and associate it with one of the two images by pressing on the possibly correct one (see, for example, Figure 7).

For the last type of exercise, Repetition of Words, two buttons are shown in the center of the screen. The first one represents the icon of a microphone and the second one that of a speaker; below, one next to the other, there will be the three words that the child will have to pronounce. By clicking on them, the image associated with the words will appear as an additional hint for the exercise. To carry out the exercise correctly, the child can reproduce the sequence of words spoken by the system by pressing the button with the speaker icon and use the hints by pressing on the words he is uncertain about; in any case, he needs to press the button with the microphone icon, pronounce the sequence of words in the correct order, and press the same button again (see Figure 8).

3.2.1 Exercise Execution

Depending on the type of exercise the child is doing, the proposed interface changes. Each time the child completes the execution of an exercise, the system will show him/her a window with a message, either written or spoken, congratulating him/her on completing the exercise, regardless of the outcome. In case the exercise is performed correctly, the message will also indicate the earned cookies, which act as a reward.
4 EVALUATION PLAN

In order to plan an evaluation of the system’s usability and medical efficiency, we involved two speech therapists from the hospital of Bari "Papa Giovanni Paolo XXIII", with whom we already collaborated in the developing process of the system. A series of meetings and discussions were carried out with speech therapists in order to get an insight into the medical aspects to consider and the appropriate features to evaluate; their expertise was fundamental during this phase. Two types of studies have been planned: Usability and Longitudinal study (Razak et al., 2010; Lazar et al., 2017). Therefore, the usability study has the goal of analysing the users’ overall experience with the system and its features, while the Longitudinal study is related to the system’s medical effectiveness and efficacy and the impact on the children’s speech impairments. Before planning the study, to obtain an insightful opinion about the system, a pilot study was conducted involving two speech therapists. The problems reported in this phase were resolved in advance to avoid evaluating an already-known faulty system.

This plan involves, in both cases, 10-13 speech therapists and 20-23 caregivers and 20-23 patients (because each child is associated with one caregiver). The group of therapists has also supported the process of drafting questionnaires, tasks choice, and the duration of the longitudinal study. In order to guarantee participants’ privacy, it has been established to assign to each of them an identification code.

4.1 Usability Study

The usability study consists of the execution of tasks and filling out questionnaires. This study is planned to be conducted in a remote way by using the platform "eGLU-BOX”, a software to support usability testing based on the eGLU LG 2018.1 protocol (Federici et al., 2021). The tasks have been chosen carefully by considering all the main functionalities of the system and the domain requisites. The creation of the testers’ accounts on eGLU-BOX will be carried out before the administration of the test in order to speed up the process. Two different studies have been created: one for speech therapists and one for caregivers, called respectively “Test iniziale e-SpeechT – Logopedisti” and “Test iniziale e-SpeechT – Caregiver”, respectively. Each task will be evaluated taking into account the stress factor that often alters the behavior of participants. Therefore, two additional steps will be taken before starting the test. The first is an explanation of the study and its goal will have to be carried out before starting the test. This evaluation plan also includes an introductory questionnaire in which each participant is asked to provide personal information in order to profile them and evaluate their experience in relation to their age, experience with technology, and systems supporting health. The second step is a video demos of e-SpeechT, in which they will be provided with a first overview of the system and its main functionalities, in order to let them gain comfort and reduce the stress factor for novice users. The demos at hand bring attention to possible malfunctions and anomalies that might be encountered, explaining how to recover from them. The maximum duration planned for each task is 5 minutes.

4.1.1 Study with Speech Therapists

This study involves 10-13 therapists and the execution of 6 tasks; for each task, the participant will have 5 minutes to carry it out. The title and the instructions for each task has a title and instructions are listed below:

1. "Create a therapy": fill out the form with the requested information and add some exercises to assign to the patient.
2. "Monitor statistics": correct and mark exercises from the therapies section for a patient.
3. "Create exercises": create an exercise for each category (Naming Images, Repetition of Words, and Minimum Pair Recognition).
4. "Create a new word": create a new word, fill out the form with the requested information and add its image.
5. "Modify thresholds": change the thresholds and make some trials to check if the system’s behavior has actually been affected.

1https://egluboxpa.ddns.net/public/
4.1.2 Study with Caregivers

The designated number for this part of the study is from 10 to 13 caregivers. Analogously to speech therapists, caregivers will be asked to carry out the following tasks:

1. "Set a password for a new child": set a password by choosing among one of the cartoon characters.
2. "Start a therapy": start the therapy for a child for which the button is enabled.
3. "Monitor statistics": view the statistics and results of the exercises carried out by a child.
4. "Stop a therapy": Stop a therapy for the child whose therapy has been previously started.

4.2 Longitudinal Study

The longitudinal study has been planned with the goal of measuring the effectiveness, the ease of familiarization with the system, and the differences between the traditional and experimental approaches of the system. In order to perform a proper evaluation, the between-subject design technique has been chosen, which consists in forming a control and an experimental group. The planning phase was carried out after an accurate literature analysis of the use of diaries and data gathering in longitudinal studies (Lazar et al., 2017); in this case, diaries are in the form of questionnaires that are filled out periodically to understand how the interaction with the system evolves over time. The recruitment of children and their caregivers is delegated to the speech therapists, who are in charge of creating two homogeneous groups in order to balance out potential individual differences that could ruin the end results. After many meetings and an in-depth discussion, it was concluded that since the smallest period of time necessary to obtain medically relevant results and to check the impacts of therapies on children is 6 weeks, the estimated duration of the study is 6-8 weeks. The effectiveness and relevancy of continuous communication with therapists during longitudinal studies can be observed in the literature, too; an example is the "Tablet-to-Tablet (T2T) Speech and Language Therapy Software Development Roadmap" (Jesus et al., 2019b), in which the testing phase of systems that support speech therapies focuses on real end-users and lasts around 6 weeks. It is worth to underline that the medical results of the study will be evaluated by the speech therapists, while computer scientists will take care of only the technical aspects.

In the following two sections, the activities of each group will be described in more detail.

4.2.1 Experimental Group

The experimental group will be formed by 10-13 patients with their caregivers and they will carry out the therapies with the support of e-SpeechT. The plan consists of the periodical administration of a questionnaire, through Google Form, that consists of questions about the main services that participants avail of, according to the type of participant. Two questionnaires have been created, and both of them contain questions about the experience of the patient in order to get an insight from two different points of view. The questionnaire for the speech therapists includes questions that concern the used device, the ease of use when creating a therapy, the automatic correction tool for exercises, and a section for possible other weaknesses or problems that emerged. The questionnaire for caregivers includes questions that concern the process of starting/stopping the therapy, the experience while flanking their child while using the system, monitoring the statistics, and a section for other miscellaneous observations.

For caregivers, the questionnaire is administered right after the child’s completion of the series of exercises planned for the day, implemented into the system. The section of the questionnaire concerning the patient’s experience remains unchanged regardless of whether it is administered to a caregiver or a therapist. They are asked about the time taken to complete exercises, the process of logging in their personal area, the emotions expressed by the child while playing and how they are adjusting to the system as time passes. Some questions are:

- "If the child encountered problems while logging into e-SpeechT, please specify them”.
- "Did the child have fun while playing?”
- "Did the child want to quit the game before its end?”

The type of response varies among Likert scale, multi-options, and open text.

Emotional reactions are crucial when it comes to evaluating children’s experiences because they are often raw and unfiltered, letting the most relevant aspects emerge. Since the target patients are children, asking them to fill out a questionnaire would be infeasible for two main reasons: the playful atmosphere would get ruined and some of them are not old enough to read and comprehend their content. Hence, caregivers and therapists will have to answer in their place. Nevertheless, a prototype of a pop-up has been designed and implemented in order to administer a rating scale question directly to the patient through a
The data obtained through these questionnaires will be analyzed at the end of the study to understand its outcomes and results; they will be compared with those obtained from the control group to draw conclusions.

4.2.2 The Control Group

The control group will be formed by 10-13 patients with their caregivers and they will carry out the therapies in a traditional way, without the help of e-SpeechT. They will not have to answer periodic questionnaires, since traditional protocols and methods do not need to be further proven as effective. The comparison will only involve the results and the impact that the therapies have had on the children.

5 CONCLUSIONS AND FUTURE WORKS

Systems that support medicine, health, and well-being are spreading more and more in today’s world. Speech therapy is a field in which technology can be extremely helpful since patients’ perseverance and time consistency is crucial for the successful outcome of treatments. e-SpeechT is a web-based system designed for speech therapists who oversee young patients with primary language disorders, followed in their therapies by a Caregiver, who can be a parent or a figure close to the child who assists her/him in the execution of therapies. Dealing with systems that are accurately designed and carefully tested is fundamental in order for them to be effective, efficient, and satisfactory.

The initial pilot study with speech therapists brought up some interesting points that highlighted the problems and strengths of the system, which resulted to be very useful to plan the usability and longitudinal studies. This study researched the most convenient and appropriate ways to evaluate e-SpeechT in question by creating an evaluation plan both for usability and longitudinal studies can be the tool to portray an analysis of the most relevant aspects of e-SpeechT. The evaluation plan also involved initial research of the potential participants to recruit. The study was not conducted right after the planning phase because of the COVID-19 pandemic. It resulted that parents and caregivers were not enticed by the idea of letting their children follow therapies on digital platforms because of the long period of forced distance learning that had just passed. Therefore, it is intended to conduct the planned studies at a later time. Gaining familiarity with new applications and systems might be demanding for some; nevertheless, the advantages that come after that first phase are highly beneficial to all parties involved. In addition, speech therapists appreciated e-SpeechT during the pilot studies and seemed to grasp its potential; they have shown interest to extend the use of the system also to children with more serious issues, such as Down syndrome.

As future works, it would be interesting to use the Medical Research Framework (Craig et al., 2008; Shahsavari et al., 2020) to understand the differences between assessing feasibility and assessing effectiveness of the system; this framework emphasizes that any intervention should be presented to the patients tailored to the context and culture of the society they live in, considering political, social, and geographical aspects (Huang et al., 2021). In conclusion, it is intended to improve the implementation of the login method for children by adding a geolocalization feature, in order to improve the system’s security and preserve users’ privacy (Baldassarre et al., 2021).

ACKNOWLEDGEMENTS

The support of University of Bari through grant “Horizon Europe Seeds 2021” is acknowledged. We would also thank the student Pier Luigi Laviano who re-engineered the system during his master thesis work, the speech therapists Isabella Roppa and Roberta Cassano from the “Giovanni XXIII” Pediatric Hospital of Bari for their collaboration in the system requirements collection, analysis, and study planning.
REFERENCES


