UNIVERSITÀ
DEGLI STUDI DI BARI
ALDO MORO

Dipartimento
di Informatica

Thesis in Computer Science

Master's degree Thesis in
Cloud Computing

Re-design and evolution of a system supporting speech therapists

Supervisor:
Prof. Antonio PICCINNO

Co-supervisor:
Prof. Stefano FERILLI

Candidate:
Pier Luigi LAVIANO

Final Examination 2019/2020

INTRODUCTION

Voice disorders occur when voice quality, pitch, and volume differ or are inadequate for an individual's age, gender, cultural background, or geographic location. These are due to inherent internal and/or external factors, such as vocal cord damage, brain damage, muscle weakness, or vocal cord paralysis that often damage the vocal folds.

A common medical situation involving different caregivers is speech therapy, which consists of a set of exercises, often in the form of games, aiming to stimulate the child’s language. Since there is not a common therapy path fitting anyone’s needs, especially due to the variety of disorders, it is necessary to allow speech therapists to create a personalized treatment for each patient.

In this work, firstly new functional requirements for Pronuntia[2] have been elicited and defined starting from Pronuntia, a prototype of an integrated remote rehabilitation system, designed for speech therapists who take care of children with language difficulties. It was proposed to assist a speech therapist with the management of his patients, creating personalized speech therapies; on the other hand,

the child, assisted by his/her caregiver, will be able to carry out the exercises assigned.

Secondly, the Pronuntia system has been completely redesigned to adapt it to the MVC pattern and make it more secure and flexible; the underline system database has been redesigned to improve its efficiency and changed the architecture to move it to the server-side; the overall system efficiency has been improved, making it more responsive to improve the user experience.

Third, the graphical interface of the system has been re-designed to improve its usability; a private section for the child has been added: in the new version the reserved area to carry out the exercises is no more shared with the caregiver. The two areas have been now separated, and a new access method for the child has been created.

Further, to carry out a screening on patients under treatment and to better cope to some speech therapists' requirements regarding a more accurate automatic correction of the exercises, machine learning techniques were analysed and implemented by integrating an automatic classification application into the system. The experimentation of this new functionality was carried out using the

voice recordings of fifteen children that allowed to train the classification model, returning discreet results that allow it to be used as a basis for future refinements.

A usability test aimed at evaluating the interaction of non-expert users of the domain with Pronuntia[2] has been performed. The participants were asked to perform the basic task involved in the interaction with the system, included the newly added features. The results demonstrate that the system is appreciated and easy to use, with a minimal graphical interface that let them focus on the tasks they must complete, and that the new features introduced were interesting and useful.

This thesis is composed as follows. The first chapter shows the history of speech therapy and the evolution in the last decades. In the second chapter is described the proposed solution, starting from the further requirements analysis phase through the re-design of the system, showing some hypothetical intermediate solutions. The third chapter reports on the implemented solution, description of some of the most relevant section of the system, together with example of the running system is given. Finally, in the last chapter the experimentation phase is described with the results.

CHAPTER 1

Speech Therapy and State of the Art

INTRODUCTION

Speech therapy is a clinical program conducted by the speech therapist as a treatment program to help children with language disorders. The speech therapist helps to provide treatment for those suffering from speech sound, resonance, voice, fluency, language, cognition and eating and swallowing disorders; they aim to implement an efficient treatment program to help solve children's speech difficulties and prevent subsequent literacy problems [1]. The speech therapist must go through some evaluations to collect information on the behavior and progress of children with speech disorders.

Assessments include understanding of children's spoken language and body language, children's expression through spoken language, production and use of sounds, their ability to use language in a social context, play, eat, drink and swallow. The approach of the treatment program includes manual therapy, which involves an object- and paper-based method; this includes a series of images to represent individual consonants, vowels, and words. Those images are presented to the children and they are asked to produce sounds or words based on the images.

The treatment program requires speech therapy sessions between the speech therapist and the children, and if ignored, may delay the acquisition of language and phonological skills. In addition, speech therapy can be very expensive in terms of money and time, and this can be a burden for parents travelling from remote areas of the city.

1.1. History of Speech Therapy

The American Speech-language and Hearing Association has defined language disorders as difficulties in "understanding and or using a spoken or written symbolic system. A disorder may refer to the form of language, i.e., phonology, morphology or syntax, the content of language, i.e. semantics, and the function of language in communication - pragmatic - or any combination of the above factors" [2]. This definition takes on a naturalist perspective, according to which a compromise is defined as a deviation from the average skill level possessed by a group of people with similar characteristics. In this sense, the definition of language impairment is useful because it covers a wide range of linguistic behavior. However, it does not help professionals to understand which

differences in language behavior constitute a compromise or according to which level of compromise a correction is required. Rather than worrying about the absolute level of language impairment, one should question the impact that language impairment has on the overall development of an individual in childhood or adolescence and the ability not to compromise the individual's normal daily activities.

The first descriptions of a language learning disorder syndrome in children date from the late eighteenth century to the early nineteenth century. Gall (1835) was one of the first to describe cases of children with difficulties in understanding and using spoken language to differentiate them from those with intellectual disabilities [3]. Subsequently, several important discoveries concerning the relationship between the brain and speech behavior in adults took place, thanks, for example, to the work of neurologists such as Broca and Wernicke. The disorders that Gall initially identified were thought to be parallel to the aphasia that these neurologists were studying in adults.

For the first hundred years after the start of studies on language learning and its disorders, neurologists dominated the field, focusing attention on the physiological substratum of language behavior.

Gesell and Amatruda (1947) were pioneers in the development of pediatrics; they proposed innovative techniques to assess language development and discovered what they later called "infantile aphasia" [4].

Between 1959 and 1964, Benton gave a detailed description of children with this syndrome and he is responsible for the more specific concept of language disorder, structured to the exclusion of other syndromes such as autism, deafness, intellectual disorders, and parallels with aphasia in adults [5]. At the same time, while these doctors were defining the notions of language disorders, another group of professionals was introducing new theories about children with language learning difficulties. Ewing, McGinnis, Kleffner, and Goldstein were all educators of deaf individuals and developed a wide range of techniques to teach language to children who could not speak or hear . They all noticed that the abilities of some deaf children were worse than expected based on their deafness alone. This observation led them to focus their interests more on the very

difficulties of language learning itself, and they decided to try their hand at developing more efficient methods for those children who could not learn with the approaches usually used to teach language to other deaf children.

McGinnis, Myklebust, and Morley supplemented the information available at the time about language disorders in deaf or aphasic children and proposed educational approaches that could be used to address the language dysfunctions these children presented [6], [7], [8]. McGinnis developed the method of association to teach language to aphasic children. This method was very influential in the development of the field, as it provided a first, comprehensive, and strongly structured approach to language teaching. McGinnis was also one of the first to distinguish between two types of language problems in children. What he called "expressive" or "motor" aphasia (what we today call expressive language disorder) and what he called "receptive", or "sensory" aphasia (what we today call receptive language disorder).

The study of language itself was revolutionized by the introduction of Chomsky's theory of transformational grammar

(1957) [9]. This innovation led to an explosion in research on language learning in children.

In the 1960s and 1970s, when child language research expanded from syntax to semantics, followed by pragmatics and finally phonology, the pathology of language followed in its footsteps, broadening the view of relevant aspects of language that needed to be described and addressed in clinical practice.

A large amount of new information gathered on normal language development made it possible for language pathologists to describe a child's language behavior in detail and made it possible to make specific comparisons with normal development of language skills, based on a wide range of forms and functions.

It has become increasingly clear from family studies and the study of twin individuals that genetic factors have a strong influence on language development and its disorders, although the discovery of a language gene is unlikely. On the other hand, it is likely that more genes, of small effect, will alter the way the brain develops, in a way targeted at small areas, but incisive, making the development path from genes to the brain extremely complex and difficult to predict.

Furthermore, nowadays it is known that children with language problems, in the absence of other syndromes, have no obvious neurological lesions that could explain their language difficulties; and that children with early brain injuries rarely have long-term deficits in language learning [10]. This assumption has led to some changes in the terminology we use to label language difficulties in children.

1.2. Speech Therapy in our days

In recent years, improvements in information and communication technologies, together with the mobile Internet, play a key role in modern health-care solutions, offering connectivity anywhere and at any time.

M-health (Mobile Health) solutions can be used in applications such as the detection and prevention of specific diseases, the decision-making process and the management of chronic conditions and emergencies, improving the quality of patient care and reducing care costs health. Mostly thanks to the advancement of Cloud Computing technologies, remote therapy solutions are currently possible and easily accessible by a wide range of people, including

children. Among the advantages linked to the use of cloud solutions, in addition to the logistical ones concerning organizational and travel flexibility, there is the possibility of observing and evaluating the patient's behavior in their ecosystem context, of responding with effective feedback to their activities, and certainly to propose innovative or experimental activities.

Various studies have been conducted to characterize the domain of semi-automatic interactive speech therapy. The objective of these studies is to assess the feasibility of a more efficient and less expensive methodology for the computerized diagnosis and treatment of individuals with neuromuscular disorders, which would support doctors in their work. This methodology includes, at its lowest level, an interactive dialogue between the individual, with compromise, and an automated system for the execution of training, the collection of patient speech, and performance feedback. The next level features a mechanism for non-experts in the domain to measure patient performance. This is done through a simple and easily reproducible scheme to label expressions, at the phonemic level, based on pronunciation accuracy. At the highest level, the speech therapist can assimilate the evaluations obtained from the students'

interactive sessions, review the facial expressions of the set of tests, provide a performance evaluation and prescribe additional therapy for the patients.

Over the last decade, a great deal of research has been conducted in the area of computer-assisted language learning and language therapy (CASLT) and research has been conducted on various aspects of voice interfaces for human-machine interaction, such as speech recognition technologies [11], audio-visual articulation training [12], and an acoustic-phonetic analyser [13].

Before language acquisition, children must learn to control their speech production with skills such as breathing, tone, intensity, and vocalization control. Developmental disabilities and neuromuscular disorders can create a delay in the acquisition of these skills, which makes children unable to begin to articulate their first sounds and first words. Speech therapists often use play activities to improve these skills.

The main step in language acquisition is the assimilation of the phonetic system of the language. Speech and language therapy, at this level, consists of sessions between the speech therapist and the patient who pronounces different words and receives an evaluation

of the correctness of the pronunciation by the therapist. The final step in language acquisition is the ability to use the language as a tool for everyday situations. When children master the phonological level of the language - articulation - they must start using this language to interact with the world. An important task is to be able to describe their environment or answer questions about it - descriptive skills - and it is also important to establish dialogues in a real environment to achieve the desired goals in their daily life, known as dialogue skills.

Traditional methods for the therapy of speech disorders are based on the direct interaction between patient and speech therapist through a series of activities developed by it for the diagnosis and treatment of the patient's disorders. This direct interaction is necessary and effective in giving personal feedback to each patient. Besides, this interaction is entirely based on the subjective evaluation of the therapist, which can vary over time, as the therapist can get used to the way the patient speaks or the patient can change the therapist for any reason.

The development of CASLT systems can overcome these two drawbacks and help speech therapists by providing a semi-automatic

speech therapy tool, which allows several patients to work simultaneously with a speech therapist, with a much more objective evaluation, whose change over time will only reflect changes in the way the patient speaks.

Many of these systems are based on Augmentative and Alternative Communication (AAC) concepts to make work more fun for patients and easier for professionals during speech therapy sessions [14]. The technologies used to improve language are ASR, a term that indicates automatic speech recognition techniques, and PV, a term that indicates techniques and tools to recognize the correct pronunciation of a certain vocabulary of terms.

1.3. Systems to support Speech Therapists

In this section some existing systems used in speech therapies have been analyzed. The variety of applications available on the market allows the use of different devices, such as tablets and smartphones, in various phases of language development and with different types of subjects. Previous research, for example, demonstrated the usefulness of the use of tablets and algorithms for automatic language analysis, to complement the traditional therapy

in presence between patient and speech therapist, already composed of manually administered exercises using photographs and images, aimed at stimulating the language of the child [15].

Systems and applications for smartphones accompanying the early stages of language development and helping the expansion of vocabulary and the development of symbolic play already exist. Their main aim is accompanying and reinforcing the acquisition of language skills in production, comprehension, and narration.

1.3.1. HAPPI SCRIVE AND KIDEWORDS

One of the most widely used activities in speech therapy to stimulate phoneme to grapheme conversion skills is to structure writing exercises with self-control, which guide the child in a path of self-awareness of error. *Happi scrive*[1] and *KidEWords*[2] are two crossword puzzle applications for children used for the development of writing skills. As in any crossword puzzle, the presence of the

[1] https://apps.apple.com/it/app/happi-scrive/id464675842 *viewed on 15/06/2020*
[2] https://apps.apple.com/it/app/kidewords-by-chocolapps/id879490139 *viewed on 15/06/2020*

predetermined boxes, corresponding to the number of graphemes of the word to be detected, acts as a self-control mechanism.

Both apps have an attractive graphical interface and the ability to define the level of difficulty. Children with difficulty in accessing the vocabulary can benefit from this game-exercise since the visual stimulus of the word to be guessed acts as a facilitator for the lexical and then spelling recovery of what is required.

1.3.2. TEACH AND TOUCH

Teach and Touch[3] is an application to support the speech therapist who accompanies children with Language Disorder in an ecosystem of games. The speech therapist can offer the child a personalized rehabilitation path on specific morphosyntactic difficulties. The wide range of games and levels, from which to choose, allows addressing the child's areas of fragility, which through challenges, improves their language skills.

[3] *http://www.teachandtouch.it/ viewed on 15/06/2020*

This game tool can be shared with the family: the latter, at home, repeats the activities carried out during treatment or performs those preset by the speech therapist, thus improving the effectiveness of the treatment.

1.3.3. TRAINING COGNITIVO

Training Cognitivo[4] is a project in which a group of Speech Therapists and Psychologists are involved in assessing and treating cognitive, language, and learning disorders, addressing their services to pre-school and school children, adolescents, and young adults. They also provide a GameCenter, a section where a child can choose a particular area to improve (which can be Reading, Writing, Language, or others) and will be offered a series of exercises in the form of games to improve in that area.

The games-exercises can consist of memory games, where the child must remember the position of two identical objects;

[4] https://www.trainingcognitivo.it/ *viewed on 24/06/2020*

completion of sentences, using images and colors; identification of the name corresponding to the object shown in an image, proposing an anagram of the letters that make up the name, and many others.

1.3.4. LOGOTOOLS

Logotools[5] is a management application for speech therapists that allows you to manage your patients. During the rehabilitation phase for each speech therapist, it is necessary to retrieve all the information about the patient, essential to make therapeutic decisions or to compare past information with the most recent ones. Often this information is shared between colleagues, external professionals, rehabilitation centers, and speech therapists at home and consequently, this makes its retrieval not easy and fast.

Logotools is useful for speech therapists to retrieve the information of patients and to help the professional in organizing

[5] https://www.logotools.it/ *viewed on 25/06/2020*

them to make access quick and easy, facilitating therapeutic procedures.

Logotools provides anamnesis tools that consists in anamnestic protocols for different pathologies of logopedic relevance to be associated with the patient during the recording. For each diagnosis, the relative anamnestic protocol is subsequently presented. Each medical history can be consulted at any time and place from your mobile device.

A second tool is Speech therapist's diary. It provides the physician with a virtual diary for each patient containing the needed information. During the rehabilitation, the therapist will be able to note down the goals set, the results of the tests administered, the activities assigned, evaluations of other professionals.

A third tool is Agenda: a calendar that allows the speech therapist to mark appointments including the possibility of periodic appointments and generating in graphic form the daily, weekly, and monthly work plan.

1.4. Pronuntia

Pronuntia is a remote logopedic rehabilitation prototype, developed for children with language difficulties. This prototype, available as a web application, aims at integrating traditional assistance and rehabilitation activities with those provided by technological supports.

Pronuntia, was designed to assist speech therapists in the management of patients and in creating personalized therapies that consist of speech exercises. In carrying out traditional therapies, speech therapists assign their patients three types of exercises: Naming Images, in which they show an image on a sheet of paper and the child must pronounce the subject represented; Minimum Pair Recognition, in which the speech therapist presents the child with a pair of images pronouncing the subject of one of the two, and the child must indicate which image it corresponds to; Repetition of Words, in which the speech therapist repeats three words aloud, and the child must repeat them correctly in the same sequence.

On the other hand, Pronuntia allows the child, supervised by an adult (the caregiver), to perform the above exercises in the form of games.

The system also allows speech therapists to create a series of exercises, or sequences of them, even of different types, to be administered to their patients at once.

CHAPTER 2

Requirements analysis and redesign of a web application to support Speech Therapists

INTRODUCTION

This chapter will describe the collection and iterative redefinition of the requirements that emerged from the interviews with the domain experts, in order to improve and introduce new features in Pronuntia[2].

2.1. Analysis and elicitation of requirements

This work consists of the redesign of prototypal system, called Pronuntia. Since the early stages of interviews, further investigations were carried out with the speech therapists that would have used the new system, which led to the modification of old requirements and the addition of new ones. It emerged the need for a system that would allow them to manage their patients efficiently, with a clear and clean organization of the graphical interface to eliminate disturbing elements, and the design of functionalities necessary for a system of this domain to be introduced. In addition, speech therapists requested to make the system's automatic correction of exercises performed by patients more heterogeneous because of the diversity of disorders from which they suffer.

2.1.1. SPEECH THERAPIST SIDE

Concerning the sections of the system related to speech therapists, they first required the system to be more streamlined also from the point of view of the graphical interface, to make the experience more direct and less confusing; in fact, using the prototype of the web application, the speech therapists realized that the playful graphical aspect involved the use of unnecessary graphic elements that make heavier the requested cognitive load; from the perspective of the children who will use the system it is acceptable to have a “cartoon” graphical style to keep high the engagement in the use, but wrong as far as the sections involving speech therapists are concerned, since they need an interface as clear and professional as possible, which allows them to use the management system avoiding any element of disturbance; for instance, since the creation and insertion of new exercises in the system is of main importance, the flow for their execution must be clear and self-explanatory.

About exercises, just as with words, the speech therapists have expressed the desire to share them with other speech therapists in the system. Currently, in fact, while words can be characterized by

public visibility (available to all users) or private (available to the individual speech therapist who creates it), the exercises can only be private. In using the prototype, the speech therapists have often felt the need to share with their colleagues some created exercises or an entire series of exercises, which proved impossible.

In the Pronuntia prototype there is a section called "Rewards", which consists of awards that the child receives at the end of the therapy. It resembles the rewards the speech therapist gives to the child during the therapy in presence, by allowing him/her to use a toy or watch a particular video. In the prototype the child obtains the possibility to watch a video of his/her interest as a reward for having successfully completed an exercise. The prototypal implementation of this section was not understood by the speech therapists: interviews with them revealed that its use was not clear, since it did not correspond to the requirements set at the beginning. They, in fact, are asked to create a "Hashtag", to be used as a keyword in the search for videos on Youtube to show to the child; this Hashtag is made known to the Caregiver from his personal page, who can then search for videos and select the most relevant ones to show as a reward for the completion of the therapy. The confusion of the Speech

Therapists begins at the basis of the whole process: in fact, as they must know the personal tastes of the child to show him videos to his liking, it is necessary that it is their Caregiver to suggest to the Speech Therapist the keyword used for the search process; then the flow of the execution goes back into the hands of the former, who specifically choose the videos that may best please the child he takes care of.

Speech therapists expressed the need for a section where they can globally control and manage the appointments made with all the caregivers at the end of the treatment of the children they are treating, just as if it were an agenda in which to mark the events. In this way, they can check or modify the appointments in a single view, without risking overlap. Also, speech therapists need a notification system to alert them when therapy for a particular child has ended, so that they can prepare and submit a new one, as well as check how they have performed the exercises and, if necessary, correct them manually.

A big limitation that afflicts the prototypal implementation of the voice recognition in Pronuntia concerns the automatic correction of the exercises: when the child performs an exercise that involves voice interaction with the system, it transcribes the voice input and,

if the word is recognized as correct within a certain predefined confidence level, fixed for all users, the exercise is evaluated as correct. Given the diversity of patients who use the system, and consequently the different pathologies they may suffer from, it was decided to introduce a mechanism through which speech therapists, thanks to their experience, can identify and customize a number of thresholds with which to "categorize" their patients, and which will be used by the system to automatically correct the exercises they will perform. Therefore, it was decided to design and integrate a Machine Learning application that would be able to classify new and old patients according to the severity of the disorder they suffer from, by exploiting the subjective evaluations made in advance on their patients.

2.1.2. CAREGIVER SIDE

A clearer distinction has been defined between the entities of the caregiver and the patient he/she is taking care of, already during the registration phase in the system; in fact, in Pronuntia[2], the basic information related to the child is entered by the treating speech

therapist, while, during his/her registration phase, the caregiver will have to enter only the information concerning him/her.

A further element to eliminate the ambiguity of a single login screen for all the users for whom the system is intended; for the speech therapist, the caregiver, and the child, in fact, there is a single login screen in which to enter their credentials to access the system, and this leads to a little distinction between the different entities. In this way, for example, if a speech therapist trying to access the system enters incorrect credentials, he receives the error message "No parent found".

2.1.3. INTEGRATION OF A MACHINE LEARNING APPLICATION FOR SCREENING AND CLASSIFICATION OF PATIENTS

During the various phases of specification of the requirements involved in the work of reengineering the web system, various Machine Learning applications potentially integrable into the system were identified following the introduction of the mechanism of customizable thresholds; these concerned the automatic identification by the system of the severity of the disorder to which each patient is affected, to bring the level of customization of the

system from the only correction of the exercises that each patient performs, to an earlier phase, in which the system automatically assigns them a category, with the relative threshold. Starting from these assumptions, two different applications have been proposed and examined.

The first proposal concerned the setting of a value, personal to each child, that would serve as a threshold for the correction of the exercises performed by him. According to this case, a patient, already registered in the prototypal version of Pronuntia, at the first access (after the integration of this application) would be asked to pronounce a series of words and/or phrases to be registered in the system, and these pronouncements would be evaluated according to more subjective criteria by the experts of the domain; the result of this evaluation would be a numerical value (potentially between 0 and 1) to be assigned to the child and defined as "patient-level", with which to correct the exercises performed.

Subsequently, an objective evaluation would have been made using a measure of the textual distance between the word to be pronounced and the transcription of the word actually pronounced by the child; measuring this distance for all the words pronounced, for

all the children used as a sample, would have obtained the dataset to be used to proceed with the experiment.

The experiment can be described as a multi-class classification problem, in which to assign to each class (e.g., Very Severe, Severe, Normal, Good, Very Good) a range of values, to which each patient should fall. By training a classifier with the data previously obtained, when a new patient arrived, the procedure described above would have been repeated until obtaining the measurement of the distances of the words to be pronounced, and it would have been sufficient to feed him this new dataset to obtain as a result the class to which the child belongs and with which to make corrections.

A negative aspect of this proposal is the high specificity of the threshold calculated by the classifier, used as a discriminant in the correction of the exercises. Since the correction system is purely automatic and objective, even if trained with measures of similarity, it could not determine exactly the degree with which the child has performed the exercise, so the assignment of such a specific threshold would certainly have led to the subsequent modification by the speech therapist to give it more tolerance.

For a second proposal, it was thought of an initial screening to be submitted to new patients of speech therapists; it would have consisted in the pronunciation of a series of words useful to detect known anomalies in pronunciation, to allow the system to automatically recognize if a patient were pathological or not.

Two different alternatives were considered for the development of the application in question.

The first hypothesis was to model a binary classifier in which each instance, related to a particular patient, would be composed of some features corresponding to different distance measurements, calculated between a specific word selected in accordance with the experts of the domain, which would be as representative as possible of any anomalies in the patient's language, and the transcription of the voice recording of the patient's pronunciation of that same word.

The alternative would have been the modeling of a binary classifier in which, unlike the previous proposal, each instance would have been composed of features that corresponded to a specific measure of distance calculated between a series of words (significant to detect anomalies in the language, chosen together with the experts)

and the transcription of the voice recording of the patient's pronunciation of those words.

In both cases, the speech therapists involved in the experiment would have labeled each patient subject of the sample as "Pathological" or "Non-Pathological", in order to train the classifier model and to allow the prediction of the class of new incoming instances, corresponding to the new patients.

Table 1 and *Table 2* summarize and describe the structure of the classifiers of these two hypotheses.

	dist1(word)	**dist2(word)**	**dist3(word)**	**Class**
Istance 1	0,2	0,4	0,6	Non-Pathological
Istance 2	0,7	0,9	0,9	Pathological
.	.	.	.	.
.	.	.	.	.

Table 1 - Initial Screening, first hypothesis

	dist(word1)	**dist(word2)**	**dist(word3)**	**Class**
Istance 1	0,2	0,1	0,5	Non-Pathological
Istance 2	0,7	0,4	0,8	Pathological
.	.	.	.	.
.	.	.	.	.

Table 2 - Initial Screening, second hypothesis

In such a context, however, the identification of the existence of pathology in a patient does not make much sense, as the status of the patient itself already implies the presence of an anomaly in the language.

2.1.4. SEPARATION OF PATIENT FROM CAREGIVER

In the initial design of the prototype, it was foreseen that a Caregiver and the Patient he takes care of had two different credentials, but this does not happen. To access their personal area to perform the exercises contained in the therapy, the patient must first log in to their Caregiver's private area, and from there, using a button, access their own. A login system of this type manifests various problems. The ideal case is when a patient logs in under the supervision of his Caregiver, who guides him to his personal play area and then leaves him free to use the device and play; other scenarios, however, would become more problematic. For example, if a Caregiver were to tell his Patient the credentials to be able to access autonomously, the latter would have the power to modify the personal information, to modify scenarios and rewards, and in general to perform actions for which the Caregiver is responsible. If

this were not the case, there would be the problem that the Patient would not be able to remember the password set up to be able to perform the access, having this, hopefully, to be composed of a combination of alphanumeric characters, with uppercase and lowercase letters or special characters, that therefore a Patient of few years would have difficulty first to remember, and secondly to write.

A serious problem found in the exercises section is the incorrect functioning of the voice recording module of the system, with which the child must interact to carry out the exercises of naming images and repetition of words. This can be due to two different causes: first, to allow voice recording the web system must have adequate authorizations, guaranteed by the presence of an SSL certificate; the second cause could be found in the algorithm developed for the capture of voice recording if it does not handle exceptional events. Due to the incorrect functioning of voice recording, the related exercises are detected by the system as incorrect, generating frustration in children who perform them - even in a correct way - and thus decreasing the general level of engagement.

In Pronuntia[2] the system communicates messages to the children in a quicker way, to prevent the willing to skip them. Also,

they are presented with a neutral sentiment to avoid feeling of frustration in the children in the case of negative outcome of the exercise.

2.1.5. USE CASES

After the analysis and subsequent elicitations of the requirements just described, a set of use cases has been developed, with the collaboration of the speech therapists of the Giovanni XXIII Children's Hospital of Bari, concerning the main activities that users perform in the interaction with the system.

Here, the actors that use the system are presented.

- Guest: visitor of the system who is not registered or logged in.
- Speech Therapist: user of the system who can create new Patients and assign them to the Caregivers. Speech Therapists can create therapies composed of exercises to administer to Patients and make appointments with Caregivers and Patients.
- Caregiver: user of the system who is responsible for the Patients assigned to him/her. He/she can create new

scenarios for the play section with which the Patients will carry out the therapies.

- Patient: user of the system who performing the exercises assigned to him/her.

The most significant use cases, corresponding to the relevant functionalities of the system, are described below.

Use Case: CaregiverSignup	
ID:	UC04
Description:	A Guest access to his Caregiver account and complete the registration
Actors:	Guest, Caregiver, Patient
Preconditions:	A Patient with an associated Caregiver must be already registered in the system
Main sequence of the events:	1. The use case begins when the User clicks on "Registrati->Caregiver" 2. While the Patient ID is not valid: a) The Guest enters the Patient ID obtained by e-mail b) The system validates the Patient ID 3. While his personal information is not valid: a) The Guest enters his personal information, i.e., name, surname, date of birth, gender, and password b) The system validates his information 4. The system redirects the Guest to the Caregiver Login page
Post-conditions:	The Caregiver's account with which the Patient ID is associated is updated with the given information
Alternative sequence of the events:	CaregiverAlreadyRegistered InformationNotValid

Use Case: PatientUpdatePassword	
ID:	UC05
Description:	A Caregiver set a new password for one of the Patient he is treating
Actors:	Caregiver, Patient
Preconditions:	The Patient account must already be registered in the system
Main sequence of the events:	Includes (Login) 1. The Caregiver chooses from his dashboard the Patient whose password must be changed, and the system shows him a page with the Patient personal information 2. The Caregiver clicks on the "Update" button 3. While the updated information is not valid: a) The Caregiver and the Patient choose the Patient's favorite image, which will be the Patient's new password b) The system validates the password 4. The system shows the Caregiver a page with the updated Patient personal information
Post-conditions:	A new password for the Patient is set
Alternative sequence of the events:	InformationNotValid

Use Case: SpeechTherapistCreateExercise	
ID:	UC06
Description:	A Speech Therapist creates a new exercise of Image Naming, Minimum Pair Recognition, Repetition of Words
Actors:	Speech Therapist
Preconditions:	At least one, two, or three words must be stored in system
Main sequence of the events:	Includes (Login) 1. The use case begins when the Speech Therapist clicks on "Esercizi->Denominazione Immagini/Coppie Minime/Ripetizione di Parole->Aggiungi nuovo esercizio" 2. While the required information is not sufficient: a. The Speech Therapist chooses one/two/three images for the exercise b. The Speech Therapist chooses the number of cookies to assign to the exercise c. The Speech Therapists chooses whether the exercise will be public or private 3. The system shows the Speech Therapist the list of all the exercises stored in the system
Post-conditions:	A new exercise is stored in the system
Alternative sequence of the events:	InformationNotValid

Use Case: SpeechTherapistCreateTherapy	
ID:	UC07
Description:	A Speech Therapist creates a therapy to assign to one of his/her Patients.
Actors:	Speech Therapist, Patient
Preconditions:	At least one Patient and one exercise must be stored in the system
Main sequence of the events:	Includes (Login) 1. The use case begins when, from his/her dashboard, the Speech Therapist clicks on a Patient's profile picture and then on "Therapies", or clicks on the therapy icon below the Patient's profile picture 2. While the information is not valid: a) The Speech Therapist enters the start and end dates for the therapy, the place for the appointment, the date and time for the appointment 3. The Speech Therapist selects the exercises to include into the therapy 4. The system asks the Speech Therapists if he/she is sure to create the therapy 5. The system shows the Speech Therapist the new therapy created
Post-conditions:	A Patient is assigned a new therapy
Alternative sequence of the events:	InformationNotValid

Use Case: SpeechTherapistSetThresholds	
ID:	UC08
Description:	A Speech Therapist update the thresholds values with which to correct the exercises for his/her Patients
Actors:	Speech Therapist
Preconditions:	None
Main sequence of the events:	Includes (Login) 1. The use case begins when the Speech Therapist clicks on "Esercizi->Threshold Settings->Update button" 2. While the values for the thresholds are not valid: a. The Speech Therapist chooses new values for the thresholds 3. The system shows the Speech Therapist the values set for the thresholds
Post-conditions:	New values for the thresholds of the Speech Therapist are set
Alternative sequence of the events:	ValuesNotValid

Use Case: CaregiverStartPatientGame	
ID:	UC09
Description:	A Caregiver give one of his/ger Patient the permission to play with the system
Actors:	Caregiver, Patient
Preconditions:	The Patient's Caregiver must have a "active" status in the system
Main sequence of the events:	Includes (Login) 1. The use case begins when the Caregiver detects the Patient he/she wants to give the permission 2. The Caregiver clicks on the "Avvia" button 3. The button text changes in "Ferma"
Post-conditions:	The associated Patient has the permission to start playing and perform the exercises in the therapy
Alternative sequence of the events:	

Use Case: PatientLogin	
ID:	UC10
Description:	A Guest access to his Patient account
Actors:	Speech Therapist, Patient
Preconditions:	A Patient with associated Caregiver and Speech Therapist must be already registered in the system
Main sequence of the events:	Includes (CaregiverStartPatientGame) 1. The use case begins when the User clicks on "Accesso->Bambino" 2. The User detects the Speech Therapist from whom he/she is being treated 3. The User detects his/her name or his/her profile picture 4. While the image chosen is incorrect: a) The User choose the right image corresponding to his/her password 5. The system redirects the logged Patient to his game homepage
Post-conditions:	The Patient has logged into the system
Alternative sequence of the events:	WrongCredentials

Use Case: PatientPerformExerciseNamingImages	
ID:	UC11
Description:	A Patient performs the Naming Images exercise assigned by his/her Speech Therapist
Actors:	Patient
Preconditions:	The Caregiver's Patient gives him/her the permission to start to play with the system
Main sequence of the events:	Includes (Login) Includes (SpeechTherapistCreateTherapy) Includes (CaregiverStartPatientGame) 1. The use case begins when the Patient clicks on the Start button 2. While the system shows the path with the exercises to perform: a. While the Patient selects an exercise not yet done: i. The system tells the Patient the kind of game/exercise he/she is going to perform and displayed it ii. The system shows the Patient the image associated to the word to pronounce iii. The Patient clicks the recording button, pronounce the word associated to the image and clicks again the recording button 3. The system shows the Patient the outcome of the exercise
Post-conditions:	The exercise is done, and the voice recording is stored
Alternative sequence of the events:	ExerciseAlreadyDone

Use Case: SpeechTherapistScreeningPatient	
ID:	UC12
Description:	A Speech Therapist perform the screening of one of his/her Patients
Actors:	Speech Therapist, Patient
Preconditions:	The Speech Therapist assigned one of his/her Patients a therapy composed of series of exercises of BVL series, and the Patients performed all the exercises
Main sequence of the events:	Includes (Login) Includes (SpeechTherapistCreateTherapy) Includes (CaregiverStartPatientGame) Includes (PatientPerformExercise) 1. The use case begins when the Speech Therapist clicks on his/her profile name, then on "Screening" 2. The Speech Therapist detect the Patient to screen for 3. The Speech Therapist clicks on the "play" button to start the screening 4. The system yells the result of the screening
Post-conditions:	The Patient is associated a new level of severity of disorder
Alternative sequence of the events:	None

2.2. Pronuntia^2 Design

Pronuntia^2, is the new system, born from the evolution and reengineering of the Pronuntia prototype. It introduces new features

that emerged from the elicitation of new requirements and old ones revised.

The first step in redesigning the web system was to adopt the MVC pattern design to delineate a separation between the entities to be modeled, the way the data is processed, and the web pages to be shown to users.

It was decided to set the server's filesystem to host all images and patients' voice recordings hierarchically. Inside the server, therefore, all the files will be stored, while, inside the database, only the references to them will be present, thus lightening the overall load and avoiding system slowdowns or paralysis. In this way, whenever there is a need to show the images corresponding to the words entered in the system by the speech therapists, it is sufficient to show the reference to the images in the database; the same thing happens when a speech therapist wants to check the exercises performed by one of his/her patients.

To mitigate as much as possible the problem of ambiguity, it was decided to create two different sections for registration - one for Speech Therapists and one for Caregivers - and three different sections for logging in - Speech Therapist, Caregiver, and Patient. In

this way, the three main entities of the system, each having their own personal and distinct credentials, can take advantage of their own subsections of the system that include the actions for which they are responsible, and avoid confusing scenarios (such as accessing the Patient's system with Caregiver credentials).

In order to manage the three different contexts of use, a navigation bar has been designed at the top of the interface that varies according to the user who is using the system and the pages in which he is navigating; in this way, for example, the Speech Therapist for each patient can view the diagnoses, manage the therapies, check appointments, and view statistics on the exercises he has done, while the Caregiver can modify personal information or manage the scenarios to be used during the exercises.

In general, in the system, to make the graphical user interface pleasant but above all usable, Bootstrap4[6] was used, a framework to create simple and clean web interfaces that can be used by any type of device.

[6] *https://getbootstrap.com/ viewed on 01/12/2020*

2.2.1. DATABASE

To take full advantage of the framework used for the implementation (discussed in Chapter 3), the database for Pronuntia[2] was analyzed and designed; in fact, it was used as a basic structure around which to model the entire system.

2.2.1.1. Homonyms, synonyms, and ambiguities

In the system we have referred to the figure of the Speech Therapist also calling him a Doctor; they are considered synonyms. It corresponds to the Italian "Medico".

In the implementation of the system, it is referred to the Patient as "child" or "user", the use of the word Patient is avoided in order not to provoke feelings of discomfort in them. It corresponds to the Italian "Paziente".

Although sometimes referring to a Caregiver's Patient the expression "of one's own child" is used, the latter does not always really correspond to the parent.

When we talk about ExerciseX, we refer to the three types of exercises present in the system, namely Naming Images, Minimum Pair Recognition and Repetition of Words. They correspond,

respectively, to the Italian "Esercizio di Denominazione Immagini", "Esercizio di Riconoscimento Coppie Minime", and "Esercizio di Repetizione di Coppie Minime".

2.2.1.2. Entity-Relationship diagram

After the design phase, the Entity-Relationship diagram was produced, described by the *Figure 1*.

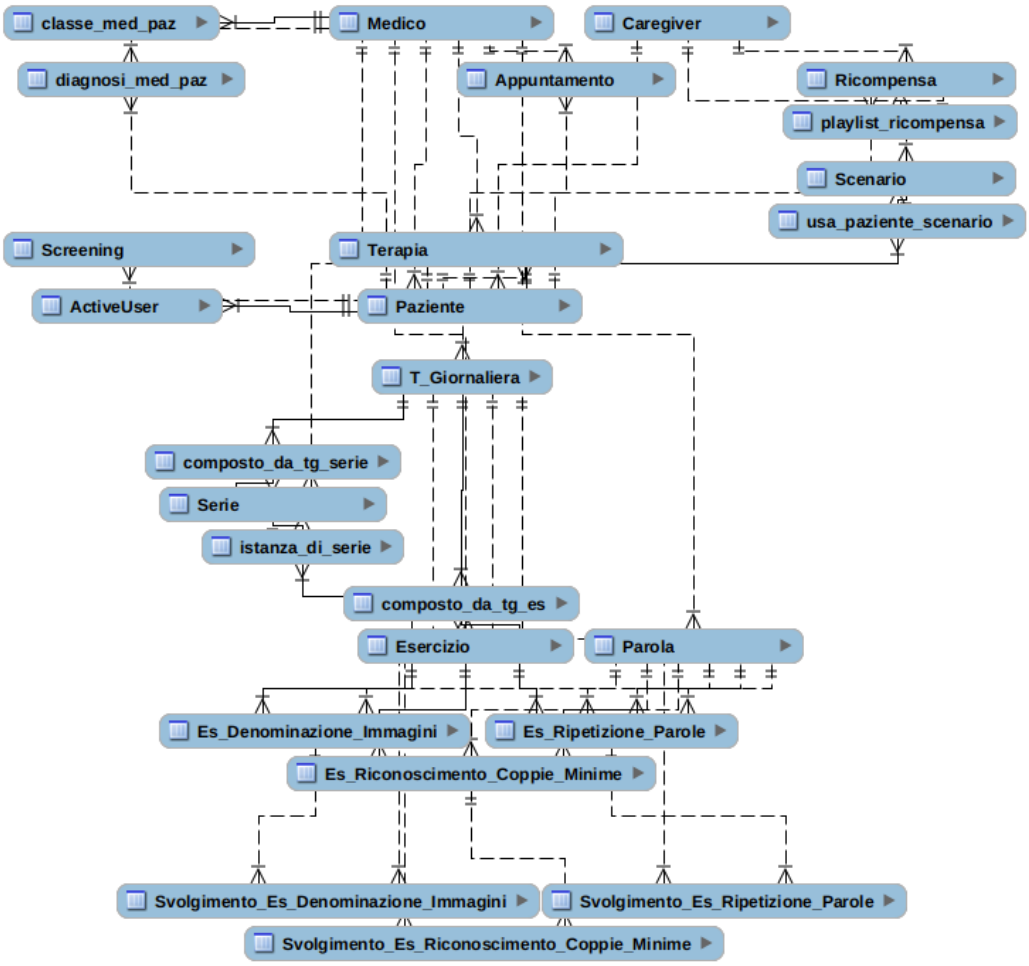

Figure 1 - Entity-Relationship diagram

It is composed by 26 tables, and the most relevant ones are described here in detail.

Entity	**Description**
ActiveUser	*Patient_id* who can start to play and *Timestamp* when he/she can start.
Appointment	*SpeechTherapist_id* who create it, *Patient_id* who is referred to, *date* and *place*.
Caregiver	*Name, surname, DOB, gender, email, password*.
ExerciseNamingImages	*Id, hint1, hint2, hint3* to help the child in the execution, *word* to be pronounced.
ExerciseMinimumPairRecognition	*Id, word1, word2, word_correct*.
ExerciseRepetitionOfWords	*Id, word1, word2* and *word3*.
Exercise	*Id* of the generic exercise, *n_cookies* i.e., the points assigned to it, *SpeechTherapist_id* who create it, *visibility*.
SpeechTherapist	*Name, surname, DOB, fiscal_code, gender, pro_pic, email, password*.
Word	*Name, img_path, description* of the subject, *SpeechTherapist_id* who create it.
Patient	*Id, password, name, surname, fiscal_code, gender, DOB, pro_pic, n_cookies* earned, *Caregiver_id, status* whether

	active or inactive in the system, *severity* of disorder.
Screening	*Id*, *patient_id* screened for, *SpeechTherapist* who perform it, *result*, *date*.
Therapy	*Id*, *start_date*, *end_date*, *SpeechTherapist* who create it, *patient_id* to whom is assigned.
Daily Therapy	*Id*, *therapy_id* it refers to, *day* of therapy.

2.2.2. COMPONENTS DIAGRAM

Pronuntia[2] has been designed according to three different components, corresponding to the actors that are involved in the system.

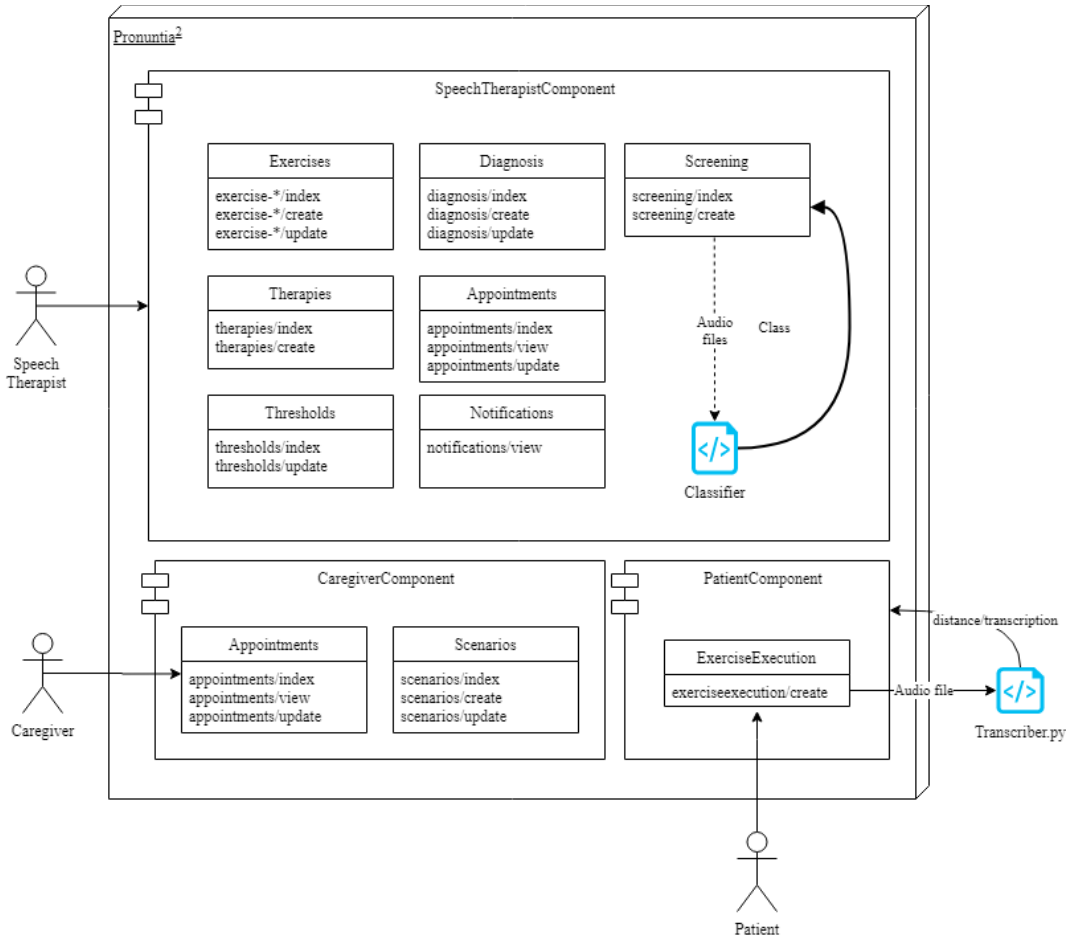

Figure 2 - Overall components' diagram

2.2.2.1. Speech Therapist Component

To give the system the most professional and minimal look possible and focus the Caregiver's attention on the relevant functionalities of Pronuntia[2], it was thought to design all (or almost all) the sections of the system in such a way as to fit all the necessary elements within a standard resolution PC browser window, without, however, obviously, sacrificing its ease of use or causing a feeling of lostness in the user. In this way, whether he is browsing using a PC, a tablet, or a smartphone, the speech therapist can immediately see all the functions he can use and all the information he has requested, avoiding having to perform one or more scrolling to reach the information of interest.

In the sections related to the words and exercises present in the system, the Speech Therapist can select the number of them to be displayed: in this way each user, also depending on the device used, can choose whether to display only a few items at a time, divided by a paging bar located below the list of words or exercises, or all together.

Concerning the Reward section, finding the double step of responsibility between the Speech Therapist and the Caregiver

inconvenient and useless, it was decided to completely relieve the former from the task of creating and managing the rewards, finding it more appropriate that it is managed entirely by the Caregiver.

To guarantee speech therapists the possibility to better manage appointments with their patients, a clear separation between the concept of Therapy and Appointment, which in the original prototype were closely linked, has been defined: since in real life the appointment does not always coincide with the last day of therapy, now the speech therapist is given the possibility to set an arbitrary date for the appointment, with the only requirement that cannot occur during the therapy but only from the last day on. Subsequently, the speech therapist can freely change this date, according to his own needs or those of the Caregiver associated with the Patient. Also, there is an "Appointments" section where, through a calendar style interface, the Speech Therapist can immediately view the appointments set daily with the Caregivers, check, or modify the details for each of them and, finally, add the event to his/her personal Google calendar, in order to integrate it with his own calendar and appointments.

A further new for the Pronuntia[2] is the presence of a Notifications section, internal in the system, in which the speech therapist is warned when one or more of his patients ends the therapy period. Through the notification, the speech therapist can directly access the details of the therapy just ended to verify the correctness of the exercises performed, or directly create a new therapy for that patient.

Finally, to obtain greater precision from the correction of the exercises involving the Patient's vocal interaction with the system - those of Naming Images and Repetition of Words - a system of customizable thresholds has been used, which correspond to the level of severity of the disorder from which each Patient suffers, within which an exercise is considered performed with success. Initially, each Patient is considered suffering from a Moderate level disorder, a value that can be changed at any time by the Speech Therapist to Severe or Slight; through the Threshold Setting section, the latter can calibrate the value for each of them, giving more or less tolerance. A lower value for a threshold translates into greater precision required by the Patient to consider the exercise performed correctly; on the contrary, a higher value for a threshold allows the Patient to obtain a

positive outcome to the correction of the exercise even with less precision. It has been decided to create a fixed number of thresholds - three - and to assign each Patient to one of these categories rather than to create a personalized value for each Patient in care at each Speech Therapist because it results more immediate than to define a specific threshold for each of them, and to avoid useless confusions in the Speech Therapists.

In the *Figure 3*, a list of modules designed for the speech therapist with a short description.

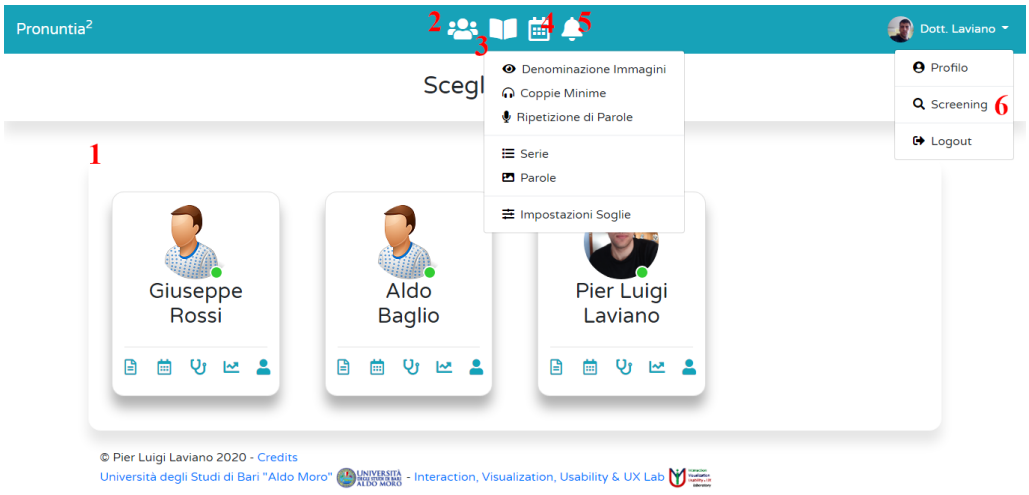

Figure 3 - Functionalities for Speech Therapist

The section labeled with the number 1 represents the Speech Therapist's view of all patients in his/her care. By clicking on the icon

labeled by number 2, he can access the section for creating a new child to add as one of his patients. From section number 3, the Speech Therapist can access the different types of exercises to be created for the system in the form of games to administer to the children; from here, he can also access the setting of the thresholds with which to allow the system to automatically correct the exercises to be carried out by the patients. The section number 4 represents the calendar with the appointments taken by the speech therapist. The section 5 concerns the personal Notification area in which the system shows the speech therapist which patients have finished the therapy. Finally, for the patients who have performed all BVL series, in the section 6 the speech therapist can perform the screening for them.

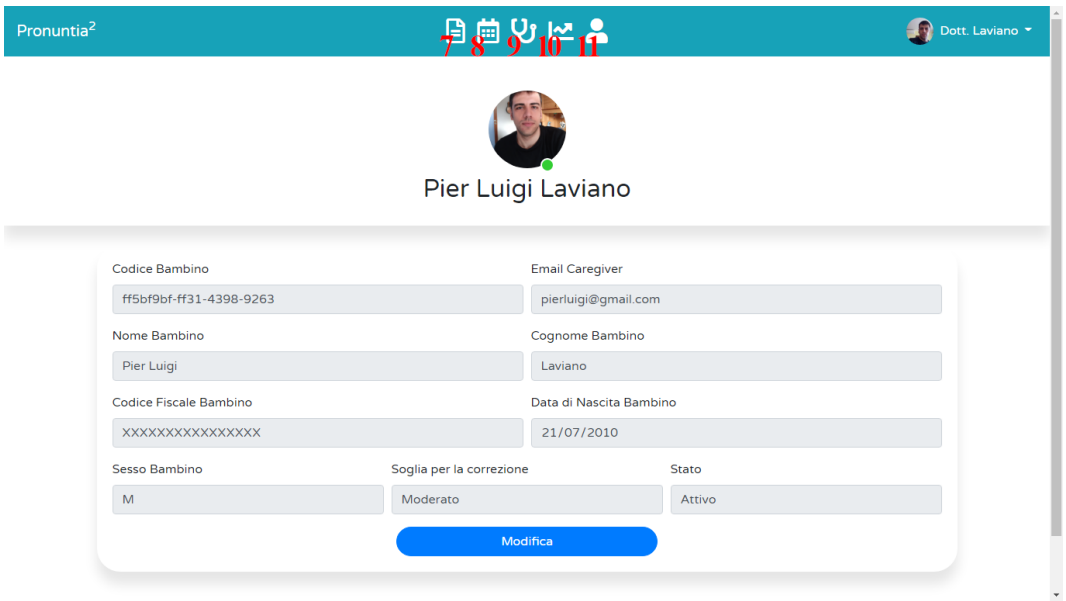

Figure 4 - Functionalities for Speech Therapist

When clicking on a patient's profile, the system context switches and the navigation bar change its behavior, showing the functionalities it offers here. Accessing the section labelled by the number 7, referring to the specific patient, the speech therapist accesses the diagnosis management in which he can view the current diagnosis, modify it, or create a new one. Through the section 8 the speech therapist accesses the calendar in which only the events concerning this specific Patient are shown. In section 9, referring to a specific Patient, the Speech Therapist views all therapies assigned to him in the past, views the current therapy if present, a scheduled therapy if it has not yet started, and can manually modify the outcome of the exercises already performed. Through the view labelled by number 10, the speech therapist can visualize the statistics concerning the outcome of the exercises performed by one of his Patient. Clicking on the icon labelled by the number 11, he can access the Patient profile in order to view his personal information and modify his current status in the system and the severity of disorder he suffers from.

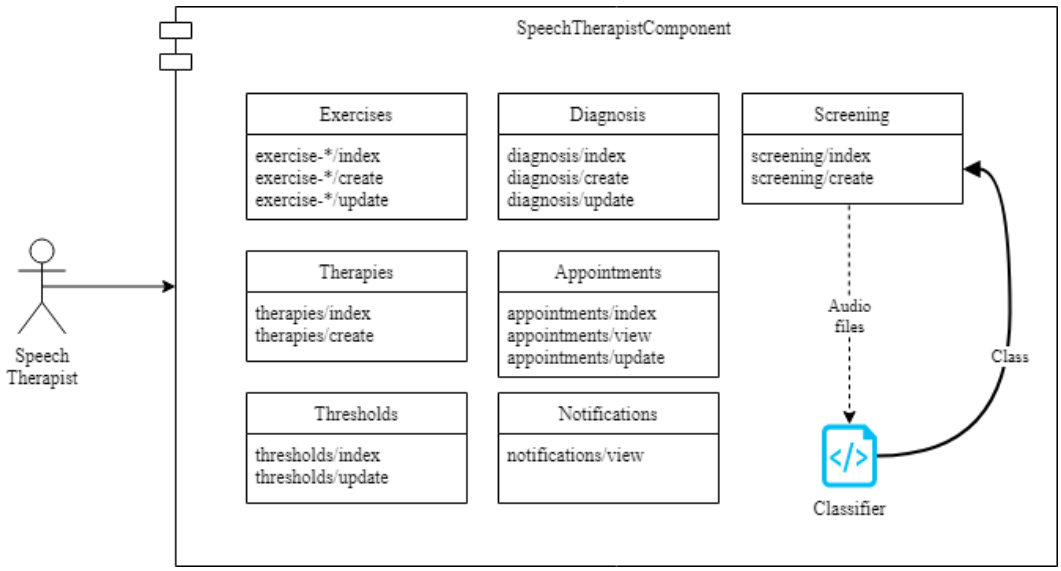

Figure 5 - Speech Therapist component

2.2.2.2. Caregiver Component

To eliminate any ambiguity and to mark the separation with the entity of the Patient, in the Caregiver registration section only the information that strictly concerns him/her are requested; the child's personal information is already entered previously by his/her speech therapist, and will then be modifiable by the Caregiver once logged in. The access now takes place from a different screen from the one used by the speech therapist.

In Pronuntia[2] the Caregiver cannot access the details of the therapy: information such as the recording of the spoken word or the possibility to manually modify the result of the exercise, in fact, fall

under the responsibility of the speech therapist only, therefore the Therapy section is not accessible from the Caregiver side.

In the Scenarios area, in addition to the insertion of an image to be used as a background for the Patient's system, it is also possible to add an image to be used as a background for the messages that the system communicates to the Patient during the interaction related to the current therapy.

Finally, the greatest introduction for the Caregiver section is the possibility to grant the Patients they are taking care of the permission to start playing; in fact, in the Caregiver's main screen there is a list of all the Patients he takes care of and, for each of them, there is a "Start" button: once selected, that Patient is granted the possibility to log in with his credentials, making his name visible among the active users in the system in the login section, in order to start performing the exercises of which the therapy is composed. On the contrary, if the Caregiver wants to prevent his/her child from starting to play or prefers to prevent it from appearing among the active users of the system, he/she can press the above button again, which now shows the label "Stop".

To summarize, the Caregiver modules of the system are in *Figure 6*.

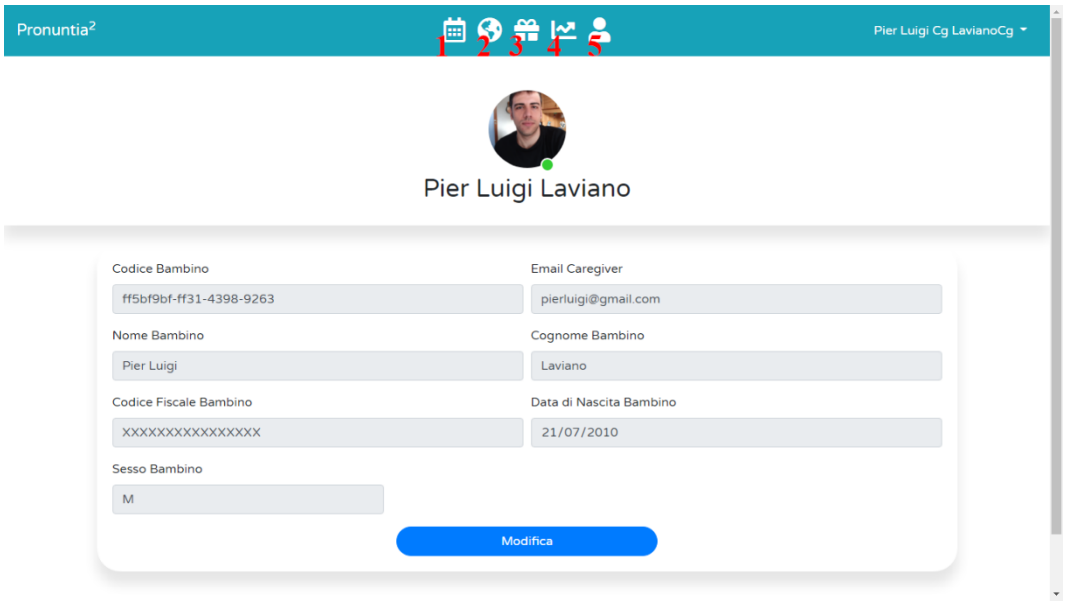

Figure 6 - Functionalities for Caregiver

When clicking on a Patient in his list, the caregiver accesses this dedicated area and the application context switches showing, in the navigation bar, the features he can use. The section labelled by the number 1 leads the Caregiver to the appointments section in which only the events taken with the specific Patient's Speech Therapist are shown. Clicking on the globe-shaped icon, the caregiver accesses the Scenario section, to customize the experience of this child. The reward section is accessible by selecting the icon number 3. As for the speech therapist, the statistics section, labelled by the icon

number 4, allow the caregiver to view the progresses that the Patient made with the therapies, in the form of pie charts. In the Patient's profile section, the Caregiver is allowed to modify his personal information, the profile picture, and the password-image he uses the log into the system.

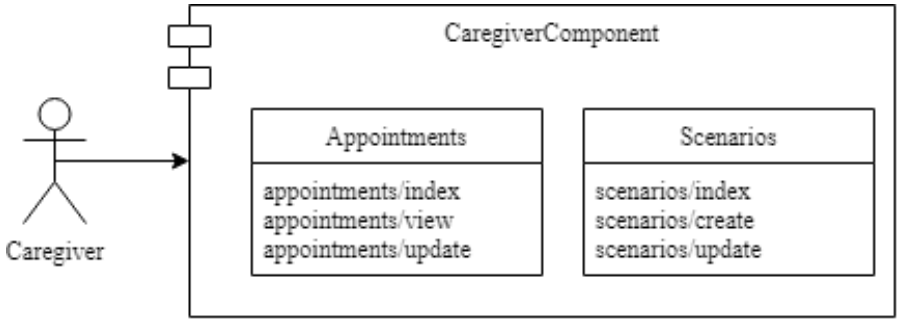

Figure 7 - Caregiver component

2.2.2.3. Patient Component

It was necessary to design a particular mechanism to allow the Patient to log in, different from the traditional method that involves the use of username and password, for more than one reason. First, we cannot expect that a child could manage safe credentials to access a system; second, even if a trivial username and password were used (in case the child can write it correctly) to allow memorability, it would run into security problems related to the ease with which it

would become possible to force them. For these reasons, it was necessary to design a system that was certainly easy for the child to use, but above all safe.

Taking the hint from the way children access the virtual classrooms of the platforms made available by code.org[7], a mechanism that would show the user the faces and/or names of the speech therapists registered in the system has been created: once the Patient recognizes and selects the Speech Therapist from whom he/she is being treated, a list of all the patients being treated by him/her is shown; the child, this time, would recognize his/her own name and/or profile image and would proceed with the last step, the insertion of a password, which consists in the selection of an image (previously chosen with the help of his/her Caregiver) that can correspond, for example, to the child's favorite cartoon or video game, which he/she will surely find easier to remember than a

[7] *https://studio.code.org/ is a nonprofit organization dedicated to expanding access to computer science education and the participation of women and underrepresented minorities.*

combination of upper and lower case alphanumeric characters and special characters.

To guarantee the first level of security, each time the order in which to show the images-password is generated randomly: in this way a possible attacker would not only have to remember in which position the password just tested is located (the first, then the second, and so on), but would have to remember the entire image. The second level of security is the one already described in the Caregiver section: from this section, in fact, a Caregiver must explicitly grant his/her child permission to access the system by selecting the "Start" button. Only in that case would the profile image and/or the child's name appear in the login section among the patients being treated by that Speech Therapist; otherwise, even if the Speech Therapist face is selected, nothing would appear, and it would be impossible to log in. Also, the Caregiver can explicitly deny permission to access the child by selecting, this time, the corresponding "Stop" button; otherwise, after thirty minutes from the granting of permission, the system automatically forbids it. Initially, it was thought to allow less time (five minutes) to increase the security threshold and limit the time window for a possible attack, but it was decided to opt for thirty

minutes to give the Caregiver and Patients with little experience of the system to have all the time available to perform the task correctly.

In the personal section of the Patient related to the execution of the exercises that make up the therapy assigned to him, the messages considered too long and/or repetitive have been eliminated and have been replaced by short and concise messages. Specifically, as soon as the Patient logs into the system, this welcomes him/her by calling him/her by name: this element may result in a higher level of engagement of the child with the system, who feels more involved and is, therefore, more inclined to perform the exercises. Once the list of daily exercises present in the therapy has been presented, the child will be able to select each one of them to begin the execution; also in this case the voice messages are of fundamental importance, because, for each different type of exercise, the system returns an introduction to the game, specifying the name and a brief description, in the form of written text and voice reproduction. This communicative paradigm has been chosen also because, being very young patients, it could stimulate the learning of reading. Finally, when the child finishes carrying out an exercise, two scenarios can occur, which correspond to its positive or negative outcome. In the

first case, the Patient is given back a message (in the same way as the other messages) with a neutral feeling that tells him/her that he/she has completed the exercise correctly, together with the number of points he/she has earned; otherwise, he/she is only told that he/she has completed the exercise correctly, thus avoiding transmitting negative feelings to the user.

During the creation of Images Naming and Repetition of Words exercises, the voice recording is captured as an audio file in base64 in a form field and is input to an external script to be processed. This Python script, Transcriber.py, has basically a double function: first, as the name suggests, given in input an audio file returns the transcription of what is pronounced within it; alternatively, in addition to returning only the pronunciation, it also returns Levenshtein's distance, used as a measure of the distance between two strings, between the word to be pronounced and the transcription of the word pronounced by the Patient. For this reason, this script can be used whenever there is a need to obtain the result of an exercise or, for example, when one only wants to obtain the transcription of the audio, in order to perform other operations on it later.

Once the audio file has been processed and the normalized Levenshtein distance has been obtained, this value is compared with the threshold defined for the child for the correction of the exercises he/she performs: in fact, if the distance value is less than the threshold value, the exercise is considered to have been performed correctly; otherwise, a negative result is returned.

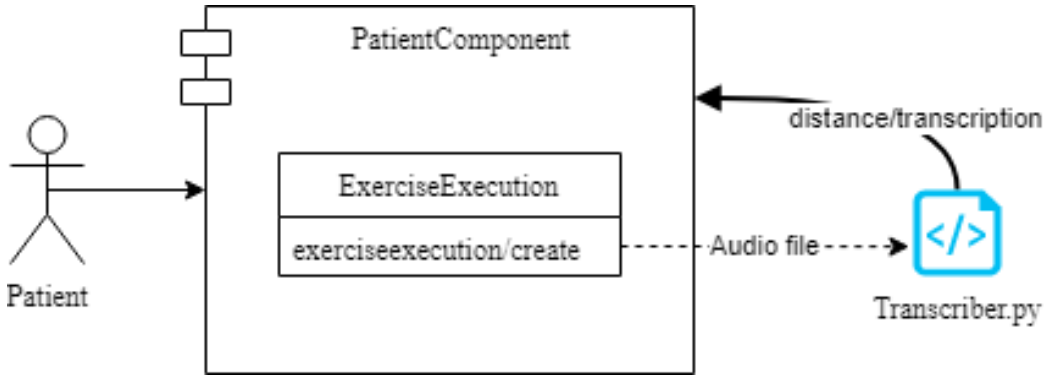

Figure 8 - Patient module

2.2.2.4. Screening and Classification Application

Using as a basis the two proposals described in the relative requirements analysis section, it was decided to design an application that takes from both the best aspects, allowing to perform a screening of patients and automatically classify them into different categories.

During the interviews with the experts of the domain it emerged the existence of a particular test, the BVL_4-12, useful to identify language disorders in children aged between 4 and 12 years, therefore suitable for the target of the system under examination.

Specifically, the BVL is divided into 3 different sections: the first section assesses oral production skills and consists of exercises in word naming, semantic and phonological fluency, and sentence completion; the second section, which assesses oral comprehension skills, offers exercises in phonological discrimination, lexical and grammatical comprehension; finally, the last section assesses oral repetition skills, i.e., the ability to repeat words, syllables, or whole sentences.

In agreement with the speech therapists, therefore, it was decided to create through the Pronuntia prototype six different series of exercises of image naming, taking inspiration from the structure of the first section of the BVL test, to be used as therapy for the sample children to be submitted to the experiment.

The objective set was, therefore, to create a system capable of allowing a patient (new or already present in the system) to perform the voice recording of the pronunciation of words in the series

previously created, and then, through Machine Learning algorithms that exploit the results obtained from the sample of the experiment, to automatically establish the class of gravity to which the child belongs; this class corresponds to one of the thresholds previously described, customized by speech therapists, according to which the system automatically corrects the exercises performed.

The application that we have chosen to develop, therefore, was born from the intersection of the two systems that emerged during the analysis phase, taking from the first the assignment of a patient to a specific class, and from the second the idea of using the application itself both as an initial screening to establish the severity of the disorder from which the child suffers, but also as an ongoing assessment of his state, as evidence of any improvement or deterioration.

It was necessary to create a sample dataset using the patients already present in the system, using the voice recordings made during the exercises. For this reason, it was asked the speech therapists to create, in the Pronuntia, six different series of image naming exercises; this subdivision of the series was made based on the phonemes present in the words to be used which, as the series

progress, become more complex to pronounce. Specifically, the first series concerns the words containing the P-B-T-D phonemes, the second the K-G-S phonemes, the third F-V-SCI, the fourth Z-CI-GI, the fifth M-N-GN, and the last L-GLI-R, for a total of 63 words.

Series name	**Phonemes involved**	**# of words**
Bvl series 1	P-B-T-D	17
Bvl series 2	K-G-S	13
Bvl series 3	F-V-SCI	8
Bvl series 4	Z-CI-GI	10
Bvl series 5	M-N-GN	6
Bvl series 6	L-GLI-R	9

Table 3 - Description of the series

At this point it was necessary to decide which measurement to use as a feature for the dataset to be generated; therefore, two different string similarity measurements were analyzed: the Jaro-Winkler distance and the Levenshtein distance.

Both measure the "modification distance" between two character sequences, i.e. how much two strings are different from each other by counting the minimum number of operations required to transform one string into the other; the difference between the two techniques lies in the number of operations available: in fact, while

in the Jaro-Winkler distance only the transposition of characters is considered, in the Levenshtein distance the operations of removal, insertion, and substitution of characters in the string are considered.

For this reason, besides the fact that it is a language-independent approach, it was decided to use the latter measure as a value to be assigned to the dataset features.

The functioning of the screening and classification application, therefore, is guaranteed by the combined use of the first Pronuntia prototype and the new Pronuntia^2, with the components already described in the previous sections. The diagram of the components of this subsystem is presented below, and for each of them a brief description will be given.

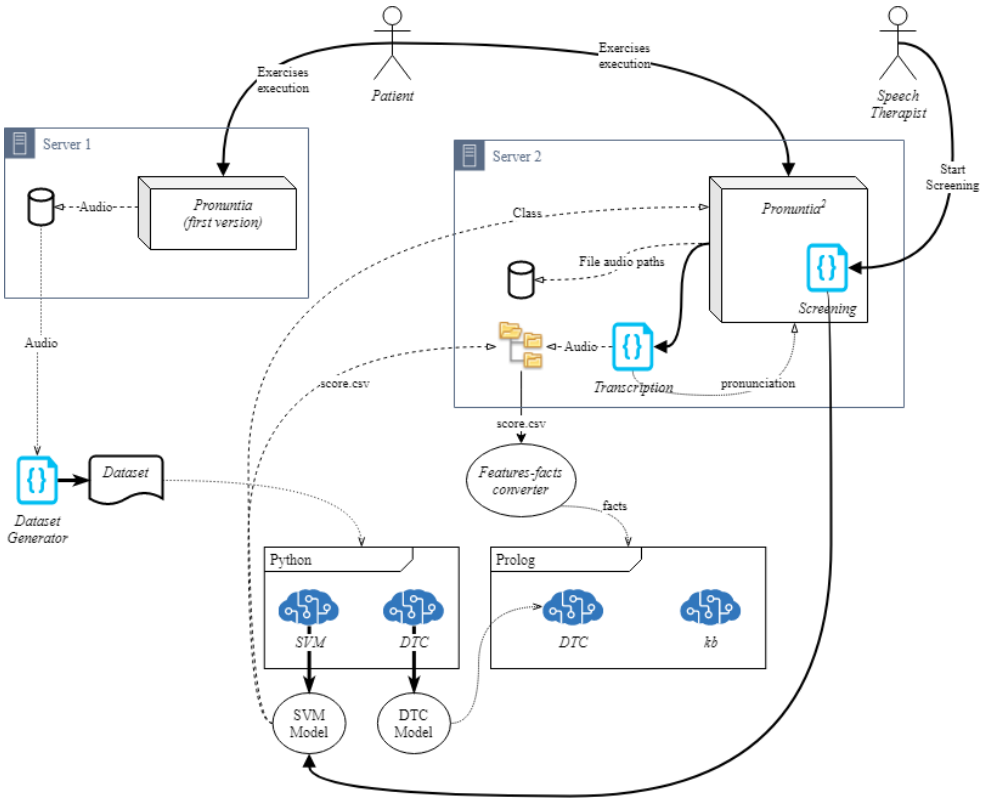

Figure 9 - Overall system with ML functionalities

The Python dataset_generator.py script is the link between the prototype and Pronuntia^2, as it allows to generate the dataset used for Machine Learning features, containing the value of the distance of all the words to be pronounced with the pronunciation performed by all the patients of the system itself. Appended to the end of this dataset, there is a last column “class”, composed of the patient’s classification carried out by the Speech Therapists. For its functioning, however, it is necessary to provide as input a csv file containing three columns, where the first corresponds to the

identification code of the patient who made the registration, the second column corresponds to the word to be pronounced, and the third corresponds to the string that represents the base64 transcription of the word pronounced; this csv can be obtained by querying the Database of Pronuntia.

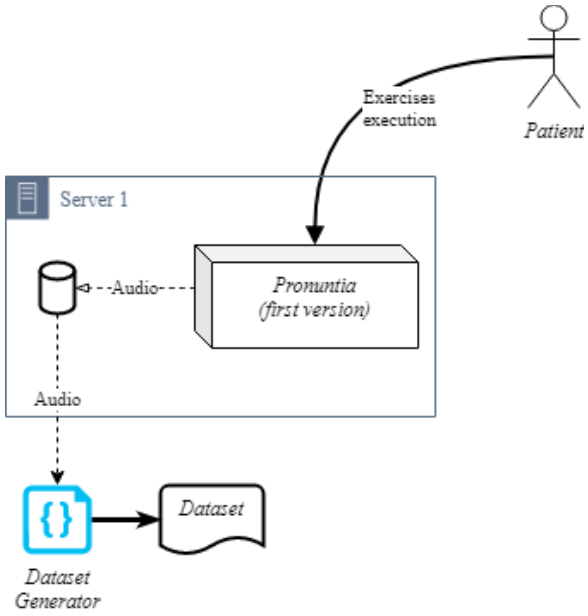

Figure 10 - Dataet Generator

For the classification problem, it was decided to opt for only three different classes to describe the severity of the patient's disorder, such as "Severe", "Moderate" and "Slight", since, in the opinion of experts, sufficient to discriminate with satisfaction against the patient population, users of a system of this type. The problem under examination can be described as a multi-class classification;

this type of technique is used when there is a classification activity involving more than two categories and each sample is assigned only one label, therefore suitable for this context.

For the design of the classifier, a Support Vector Machine Classifier was modeled in Python, which provides useful libraries, such as *scikit-learn*, to perform predictive data analysis. The dataset used is the one containing, for each instance, the Levenshtein distance value of the word pronounced with the one to be pronounced, and its classification. This file was imported into Python through the use of a dataframe, a structure made available by the *pandas* library that allows you to easily model table datasets of this type.

For experimentation purpose, a Decision Tree Classifier was implemented too, using the same dataset as the previous one; this was made to test out whether it performed worst or better than the SVM, and to extract the classification rules and implement it for a third classifier, written in Prolog.

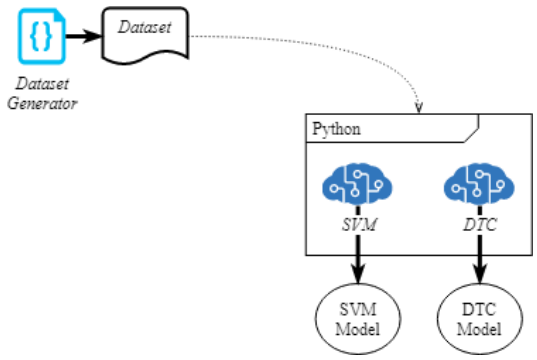

Figure 11 – Classifier

Following the flow of for the execution of the screening of a patient, a speech therapist creates in Pronuntia[2] a therapy for the latter, which includes the execution of test exercises, and assigns it to him/her; subsequently, the patient will have to perform the assigned exercises. Once all the exercises have been completed, in Pronuntia[2] there is the Screening section, accessible only to the Speech Therapist, where the latter can view the patients who have performed all the exercises related to the series of interest; these patients are the candidates for the screening, which is carried out by recalling the classification module presented above, passing in input the identifiers of the patient and the exercises performed, to be able to derive, from these, the audio files corresponding to the pronunciations made during the exercises and to elaborate the

Levenshtein distances with respect to the words to be pronounced, in order to generate a test dataset composed of these features.

The output of this module is the class predicted by the classifier, which can be confirmed and/or modified at any time by the Speech Therapist.

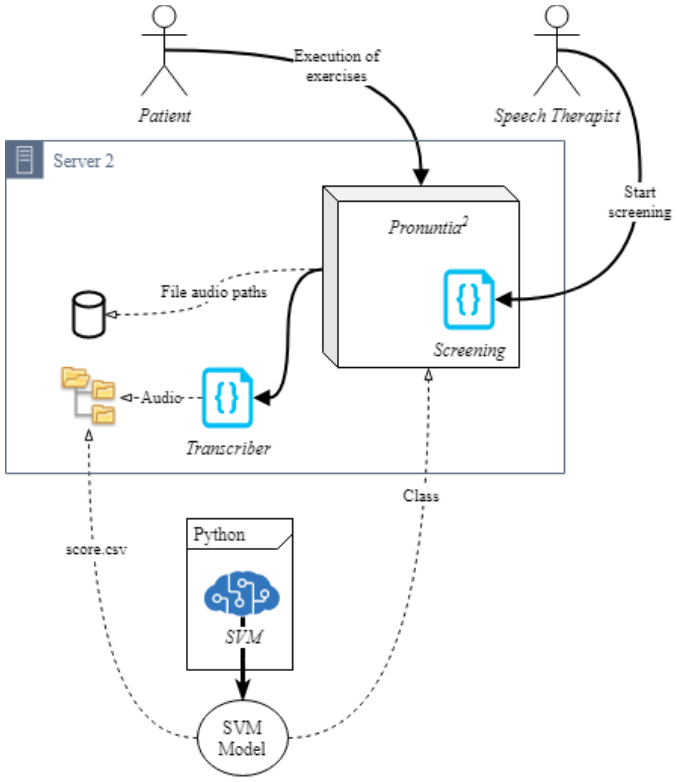

Figure 12 - Screening

CHAPTER 3

Implementation of Pronuntia^2

INTRODUCTION

Pronuntia[2] is implemented as a web application that can be used by different types of platforms (desktop PCs, smartphones, tablets) independent from the operating system, to allow users to use the device they prefer. This has been possible thanks to the use of cloud services and libraries to manage, for example, the recording of the voice input of the Patients during the exercises or the transcription of the same, with the following processing of the pronunciations to obtain the outcome of the exercises.

3.1. Technologies used

In the implementation phase technologies related to web development and scripting languages, together with tools to allow the development and testing of the entire system were used. They are better described in the following.

3.1.1. BOOTSTRAP4

Bootstrap is a free and open-source CSS framework aimed at reactive and mobile-first front-end web development. It contains

CSS-based design templates and (optionally) JavaScript for typography, forms, buttons, navigation, and other interface components.

When added to a project, Bootstrap provides basic style definitions for all HTML elements. The result is a consistent look and feel for prose, tables, and form elements through web browsers. Also, developers can leverage the CSS classes defined in Bootstrap to further customize the appearance of their content.

Bootstrap also comes with several JavaScript components in the form of jQuery plugins. They provide additional user interface elements such as dialogs, tooltips, and carousels. Each Bootstrap component consists of an HTML structure, CSS statements and, in some cases, accompanying JavaScript code. They also extend the functionality of some existing interface elements, including for example an auto-completion function for input fields.

3.1.2. MVC PATTERN

It stands for "Model-View-Controller". MVC is an application design model consisting of three interconnected parts. They include

the model (data), the view (user interface) and the controller (processes that manage inputs).

The MVC model is commonly used for the development of modern user interfaces. It provides the essential parts for the design of a desktop or mobile program, as well as for web applications. It works well with object-oriented programming, since different models, views and controllers can be treated as objects and reused within an application.

Models represent knowledge. A model could be a single object (rather uninteresting), or it could be some structure of objects.

A view is a (visual) representation of its model. It would ordinarily highlight certain attributes of the model and suppress others. It is thus acting as a presentation filter.

A controller is the link between a user and the system. It provides the user with input by arranging for relevant views to present themselves in appropriate places on the screen. It provides means for user output by presenting the user with menus or other means of giving commands and data. The controller receives such user output, translates it into the appropriate messages and passes these messages on to one or more of the views.

3.1.3. YII2 FRAMEWORK

The *Yii framework*[8] is an open-source PHP framework for modern and rapidly developing web applications. It is built around the composite Model-View-Controller model.

Yii provides secure and professional features to create robust projects quickly. It has a component-based architecture and solid full caching support. So, it is suitable for building all kinds of Web applications: forums, portals, content management systems, RESTful services, e-commerce websites and so on. It also has a code generation tool called Gii that includes the full CRUD interface creator.

3.1.4. RECAS VIRTUAL MACHINES

The ReCaS-Bari data center has been realized by the University of Bari "Aldo Moro" and the National Institute of Nuclear Physics

[8] https://www.yiiframework.com/

(INFN) within the ReCaS project (PON Research and Competitiveness 2007-2013, Notice 254/Ric). It was completed in July 2015 and inaugurated (poster) on July 9, 2015.

The ReCaS-Bari data center hosts a small cluster dedicated to High Performance Computing (HPC) to run applications that require many processing units working in parallel continuously exchanging information with each other.

The HPC cluster consists of 20 servers, each with 40 cores, for a total of 800 cores. Each server is equipped with a graphics accelerator (NVIDIA K40) particularly useful to accelerate the execution of some calculations. All servers are interconnected, with low latency InfiniBand technology, to increase the speed of data exchange between them.

The new data center ReCaS-Bari inherits from the previous Bc^2S a strong orientation to the use of open-source software: in fact, the entire data center management system is based on open-source software.

3.2. Implementation

In this paragraph the implemented solution[9] will be shown divided according to the roles covered by the different users of the system, such as Speech Therapist, Caregiver, and Patient. The system is available at https://pronuntia2.ddns.net/

3.2.1. HOMEPAGE

The main screen of the system is the homepage, which shows the user the three access options according to his role. Of fundamental importance throughout the system is the navigation bar, located at the top: it will often change its appearance during the use of the system, showing, depending on the user who is using it, the different features that can be used depending on the current context. On the home screen, it will show, in addition to the system name at the top left, the options for registering with the system or logging in.

[9] *The website can be found at https://pronuntia2.ddns.net/*

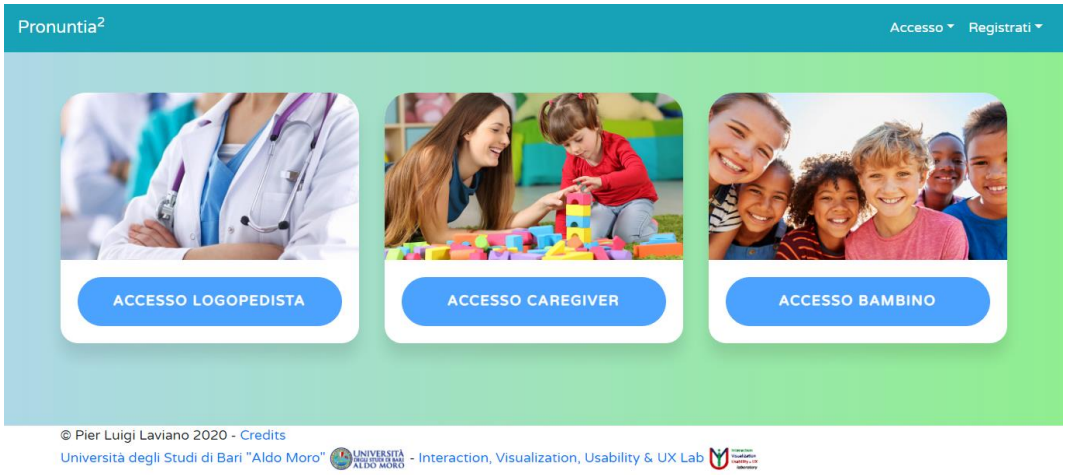

Figure 13 - Homepage of Pronuntia²

Accessing the "Signup" section, the user must choose to register as a Speech Therapist or Caregiver: if the former is selected, the system will prompt the registration form to fill with his/her personal information; if the user who wants to perform the registration is a Caregiver, he/she must indicate the code of the Patient he/she takes care of and then, if he/she is not yet registered into the system, must enter his/her personal information as well.

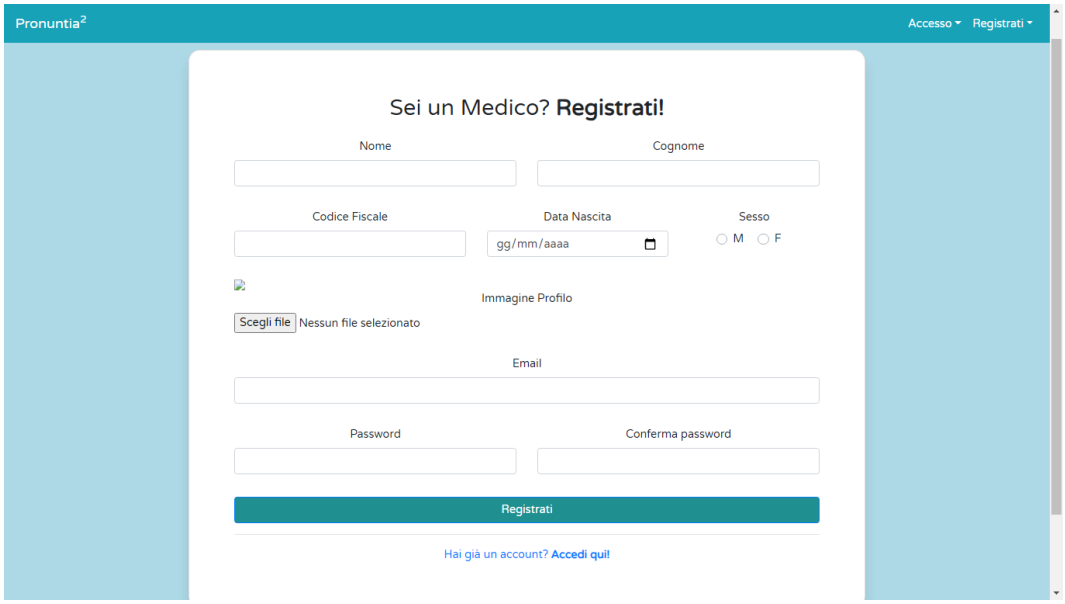

Figure 14 - Speech Therapist registration form

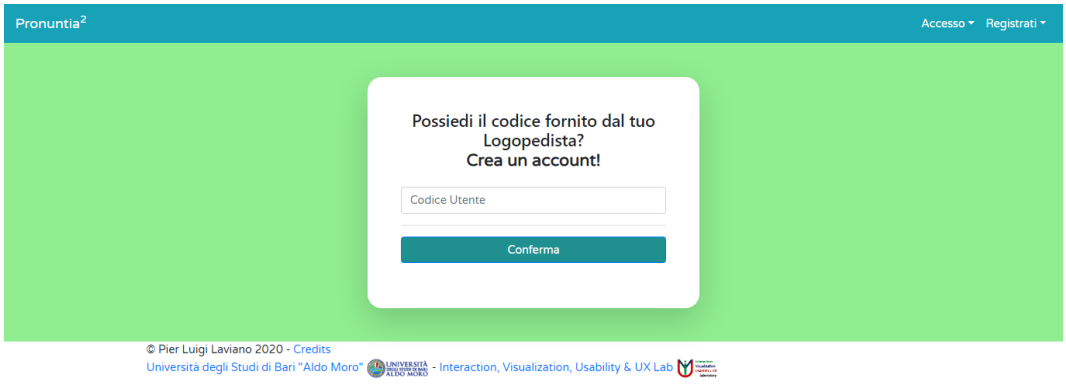

Figure 15 - Caregiver registration - first step

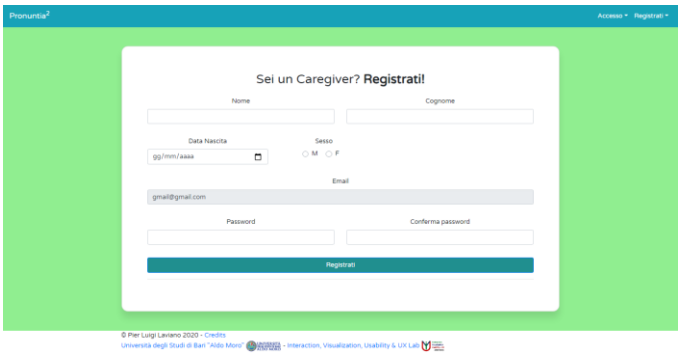

Figure 16 - Caregiver registration form

3.2.2. SPEECH THERAPIST

When the user logged in to the system is a speech therapist, he will be shown his personal main page with a list of all the children he has in charge. Here the navigation bar has changed its appearance: in fact, it shows the Speech Therapist the actions he can perform with the system at a general level, operations that do not involve changes to a specific Patient but only modify his experience with the system.

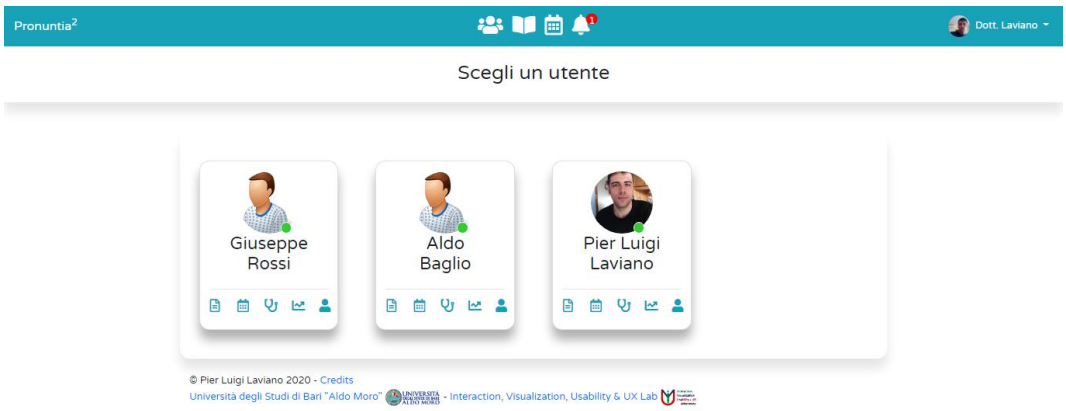

Figure 17 - Speech Therapist main page

3.2.2.1. Exercises section

The most interesting section for the Speech Therapist is the one relative to the exercises, that is composed of the management of the words, the three different types of exercises, and the thresholds settings.

The index page for the management of the words shows the list of all the words present in the system that the Speech Therapist is allowed to view, i.e., the ones he/she created or that made public by other Speech Therapists.

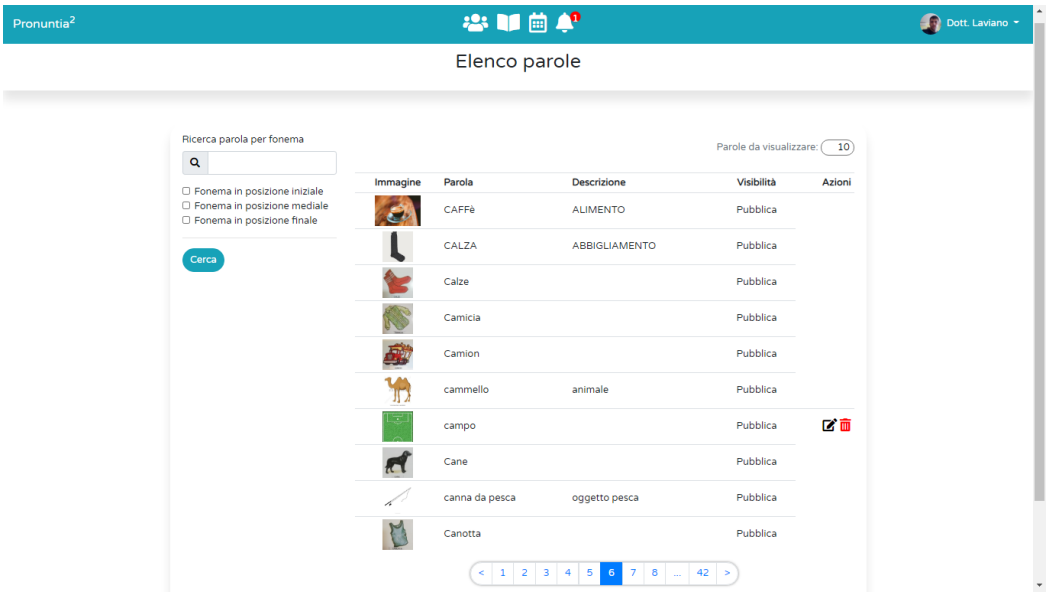

Figure 18 - Words index

Similar screens are shown the Speech Therapists in the sections concerning the index of all the Exercises present into the systems; these pages let him to filter out specific words/exercises he wants to check if present through the search form on the left side; on top of the list the Speech Therapist can choose how many elements he

wants to see at a time, and the pagination bar below the list let him change the page of interest.

For the creation of the three different types of exercises, a compact design was chosen; in fact, regardless of how many words the user has to enter, everything will fit into one screen, without the need to scroll down the page.

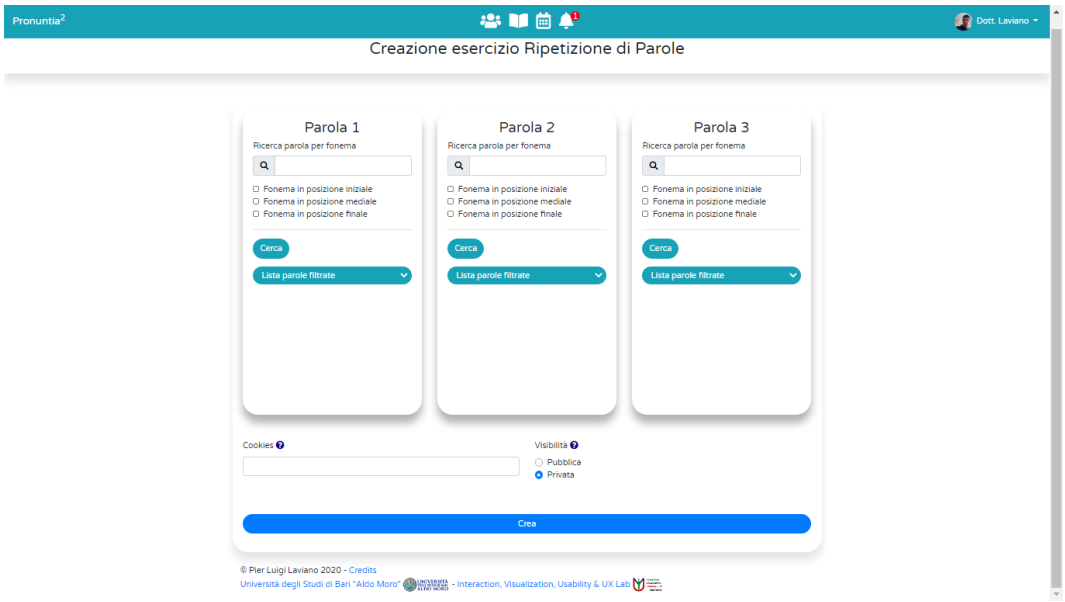

Figure 19 - Repetition of Words, that involves the insertion of three words

3.2.2.2. Patients

If the Speech Therapist clicks on one of his Patient, the system context now is focused on the latter: in a similar way as happened before, even at this point the navigation bar changes its behavior,

showing the Speech Therapist the operation that, now, will affect this particular Patient.

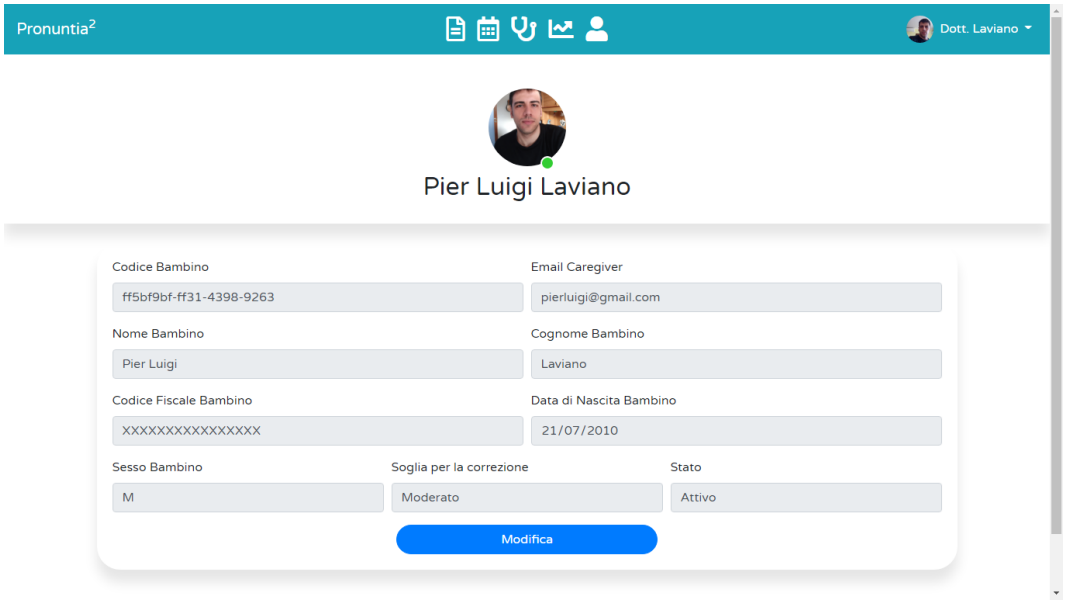

Figure 20 - Patient overview

From this section, the Speech Therapist, following the icon link present in the navigation bar, for this Patient can access the diagnosis he made, the Appointments taken with the Patient's Caregiver, can manage the therapies he created, can view the statistics about the execution of the exercises and change the severity level of disorder.

3.2.2.3. *Diagnosis*

From this section, relative to a specific Patient, the Speech Therapist can check the diagnosis' history and create a new one, at the end of a Therapy.

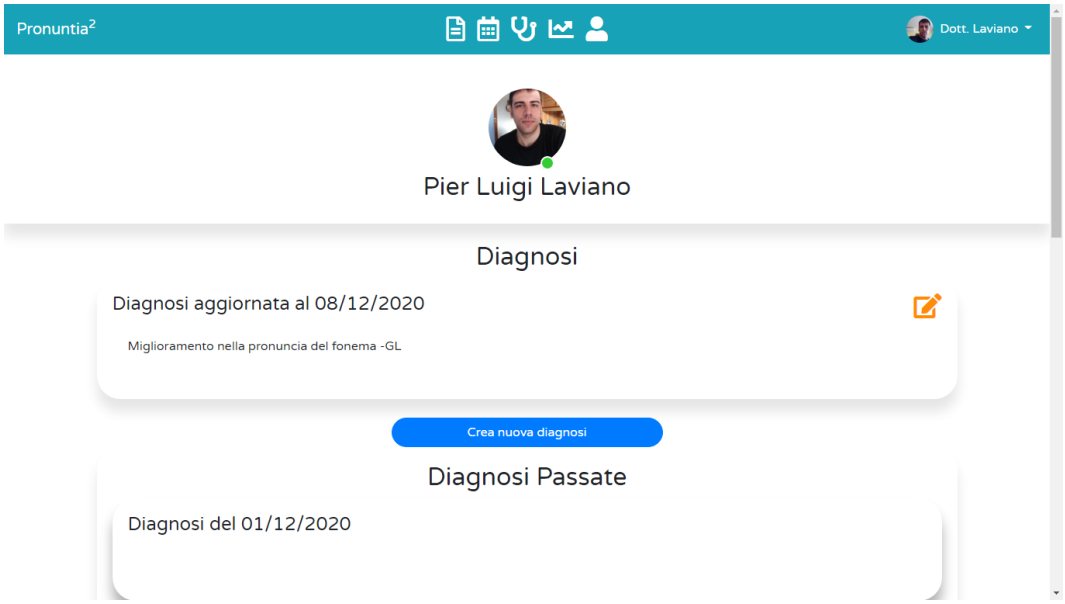

Figure 21 - Diagnosis

3.2.2.4. *Appointments*

In the Appointments view, the Speech Therapist can check the appointments taken with the Caregiver of each Patient who followed a Therapy, seeing also the details of the event such as the place and the date and time, and he can also add it to his personal Google calendar through the "plus" button.

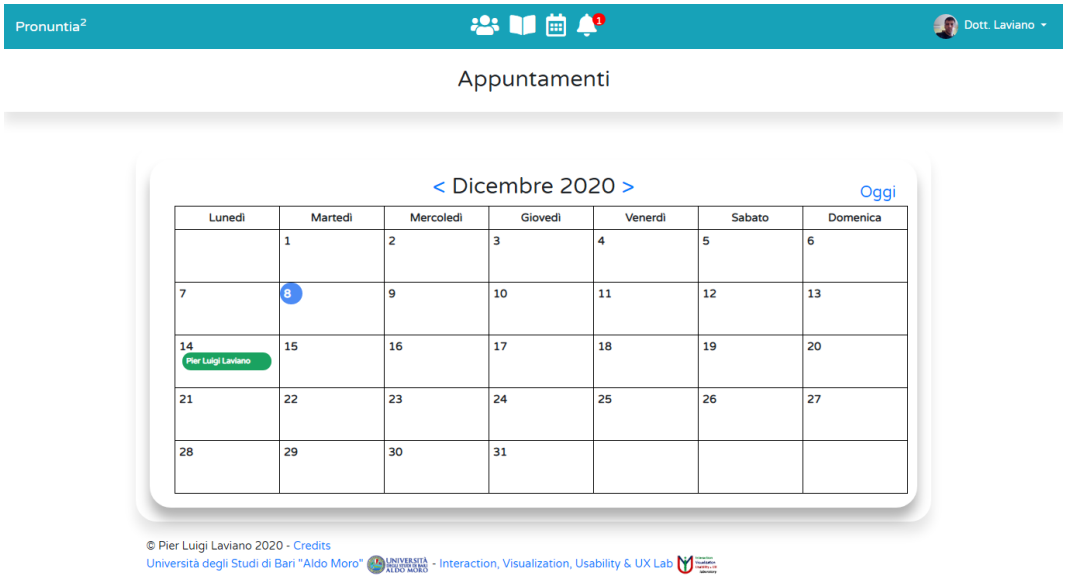

Figure 22 - Appointments view

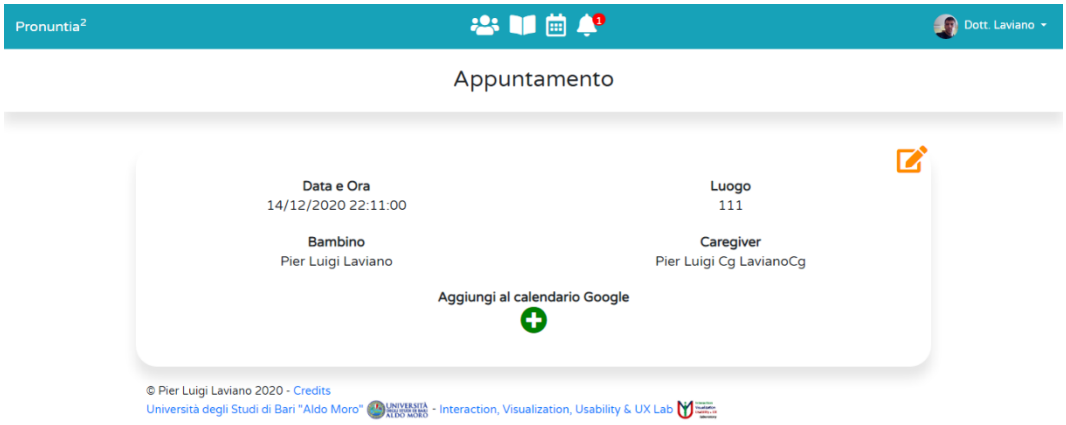

Figure 23 - Event details

3.2.2.5. Therapies

The Speech Therapist, once he selects a specific Patient, can create a new Therapy for him – if none already in progress – or he can view the details of past Therapies and future ones.

When he checks the details of a past Therapy, he can choose one of the exercises performed by the Patient to see the related details, i.e., the date of execution and the Patient's name and, for each kind of exercise, the specific information it concerns (e.g., for a Naming Images exercises, he will see the word the Patient had to pronounce, the hints he used, and the voice recording performed).

For all the exercise, the Speech Therapist can also see the outcome, and, if necessary, he can modify it.

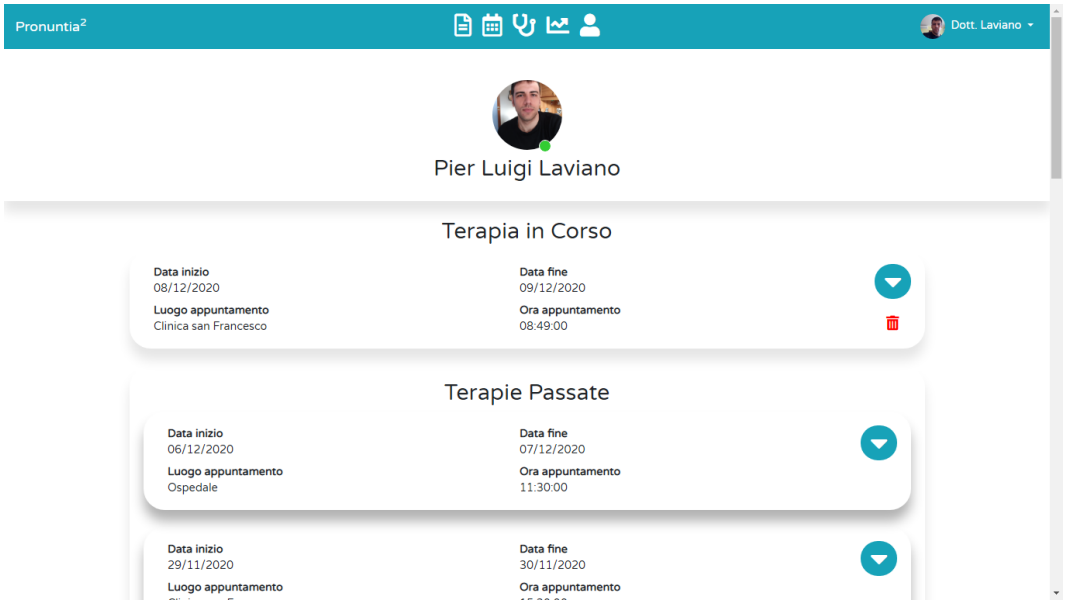

Figure 24 - Therapy section

Figure 25 - Exercise details

3.2.2.6. Statistics

When a Patient is selected, the Speech Therapist can check the screen with statistics about the exercises he has performed during the assigned therapies. The statistics will be displayed in the form of pie charts, divided according to the type of exercises performed, plus an overall one that includes the performance of all types of exercises. When the Speech Therapist hover the mouse pointer over a side of the chart (or, in the case of a smartphone/tablet, he taps on it), it will return the number of exercises performed correctly/erroneously.

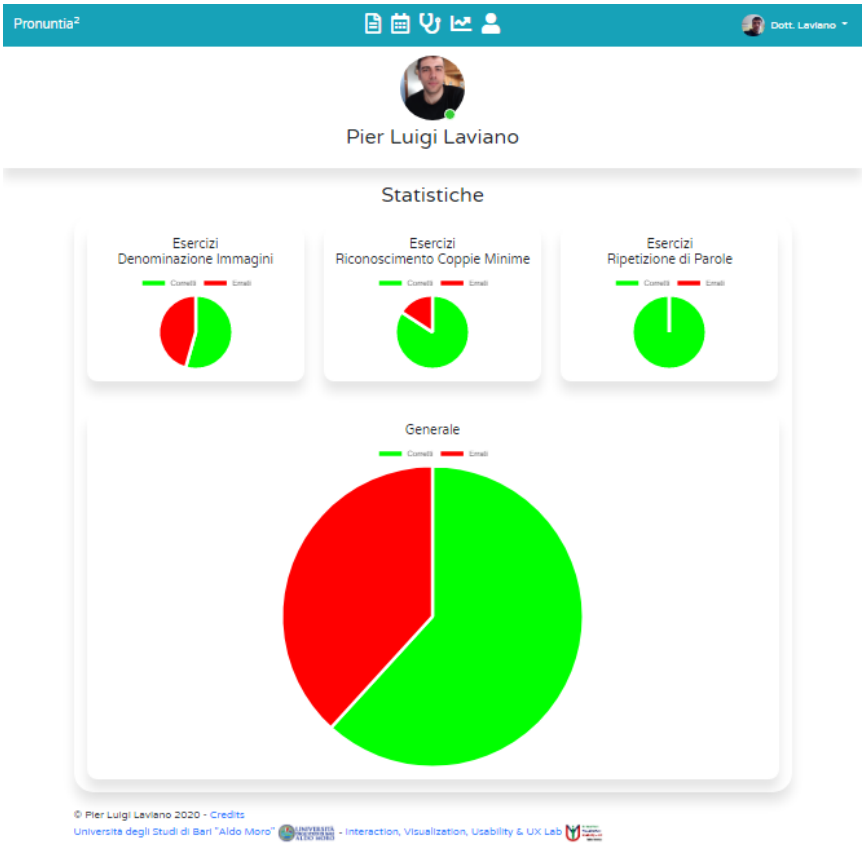

Figure 26 - Statistics view

3.2.2.7. Screening

The last main section available for the Speech Therapist is the one concerning the Screening application; here, the system will show him the list of all the Patients who performed the BVL series and, for each of them, he can start the classification. The system will return the severity of the disorder it predicts the Patient is affected of, and the Speech Therapist, if not satisfied, can modify it, fixing the classification model.

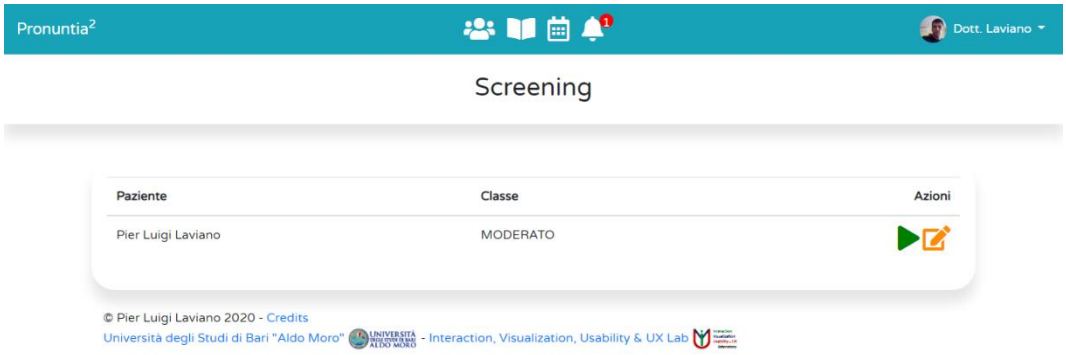

Figure 27 - Screening application

3.2.3. CAREGIVER

The main page for the Caregiver is very similar to the one showed to the Speech Therapist; in fact, while for the latter the list of all his Patients in charge is shown, here, the former, can view the

list of all the Patient he takes care of. In the main section the navigation bar for the Caregiver is quite empty, because, differently from the Speech Therapist, he does not need to customize his personal experience with the system; on the contrary, he must personalize the experience for each of the Patients he takes care of, through the Scenarios and Rewards sections. Form this main screen, the Caregiver can also give his Patients the permission to start to play with the exercise in the Therapy, if there is any in progress.

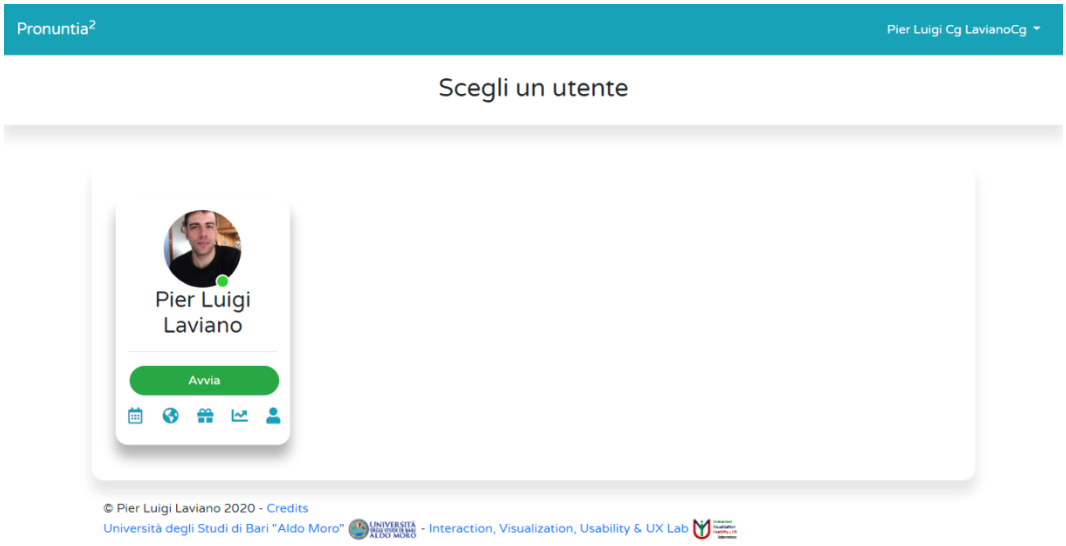

Figure 28 - Caregiver main page

If this Patient was just added by his Speech Therapist, the Caregiver must choose a new password for him, to let him play with

the exercises of the Therapy assigned. In order to do this, the Caregiver is guided by the system, that leads him to the profile editing section for that Patient and it gives him the instructions to set a new password-image for the child.

Figure 29 – Patient's password settings

In this page, that is the same from which the Speech Therapist can change the Patient's active status and level of the severity of disorder, the Caregiver is allowed to edit the Patient's personal information, such as name, surname, date of birth, fiscal code or profile picture.

3.2.3.1. *Patient management*

When a Patient is selected, the system switches the application context changing the navigation bar's appearance and behavior: at this point, it will show the Caregiver the functionalities he can access to, i.e., the Appointments section (same as for the Speech Therapist), Scenarios section, Rewards, Statistics (same as for the Speech Therapist) and the Patient's profile.

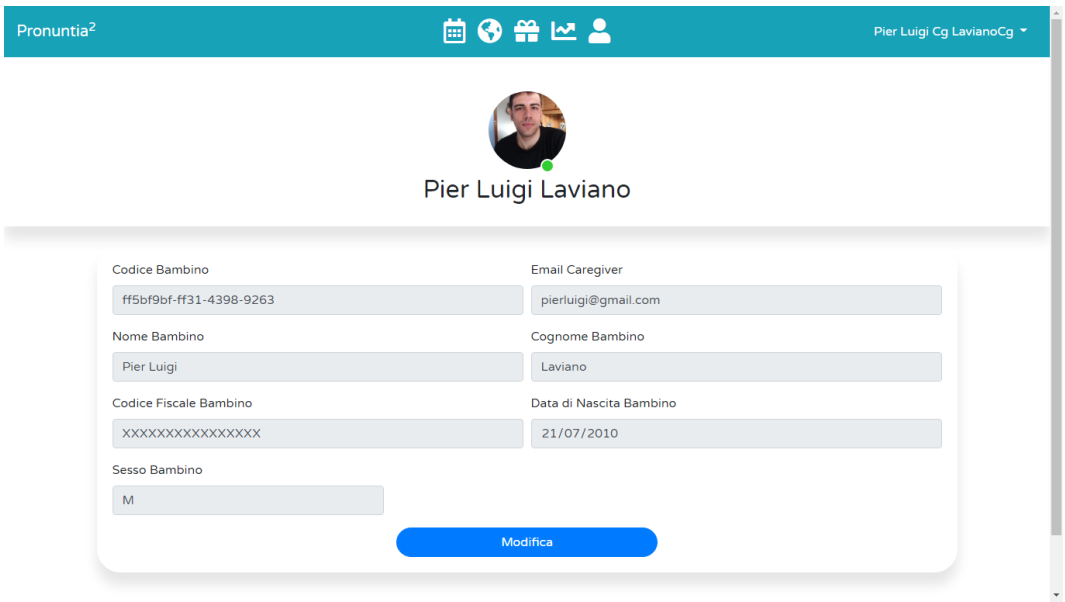

Figure 30 - Caregiver's Patient view

3.2.3.2. Scenarios

The Caregiver side section that requires the greatest involvement and customization by the Patient and him, is the one related to the Scenarios. The Scenario, in fact, defines the graphic style that the area dedicated to the Patient's game will assume during the therapy sessions. Each scenario in Pronuntia^2 is basically composed of two images: the first is the background of the system and will cover the whole screen of the section; the second, instead, consists of the background to be used for the messages that the system communicates to the child when it must indicate the instructions for carrying out an exercise, to communicate the result, or to warn him that an exercise has already been completed.

In this section, therefore, a good collaboration between the Caregiver and the Patient can result in a greater active involvement of the child in the use of the system, who will be encouraged to use it by the graphic style familiar to him/her.

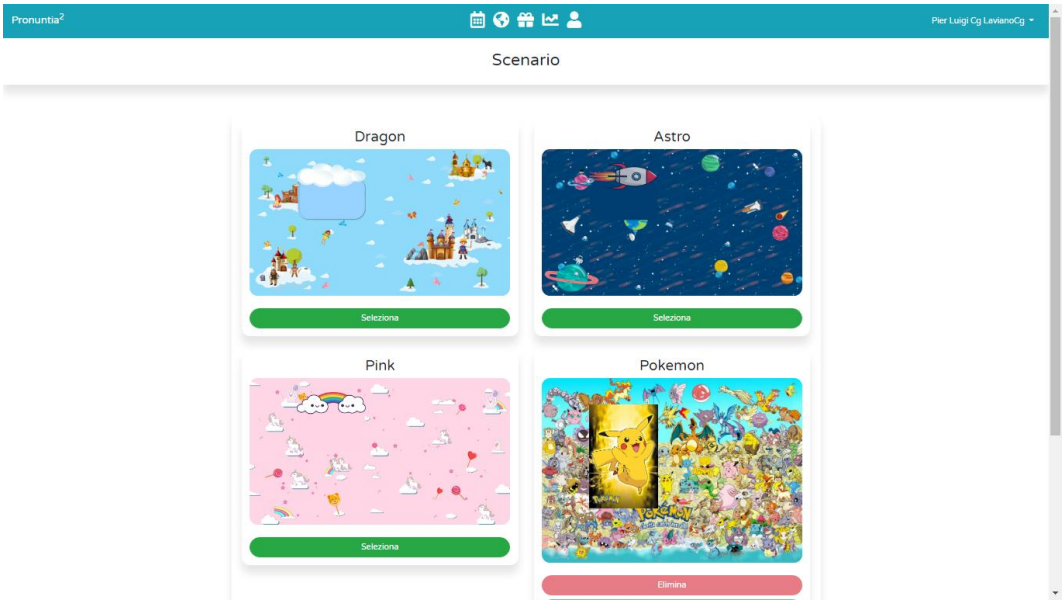

Figure 31 - Scenarios section

3.2.4. PATIENT

The Patient, as already mentioned in the previous sections of this thesis work, in Pronuntia^2 can autonomously access his dedicated play area in the system using a separate access way from the Caregiver's one; the new login method has been designed to ensure a clearer separation between the two entities, while still providing the patient with an easy way to access the system and continuing to ensure a high degree of security.

3.2.4.1. Login method

This is designed by showing the Patient, once his access method has been selected, the list of Speech Therapists in the system with their profile images; once his doctor has been recognized, the Patient can select him and, if he has been granted permission to access his play area by his Caregiver, he will be shown his name with the associated profile image. By selecting his identity, he will only need to select his password-image chosen together with his Caregiver to finally be able to access the system.

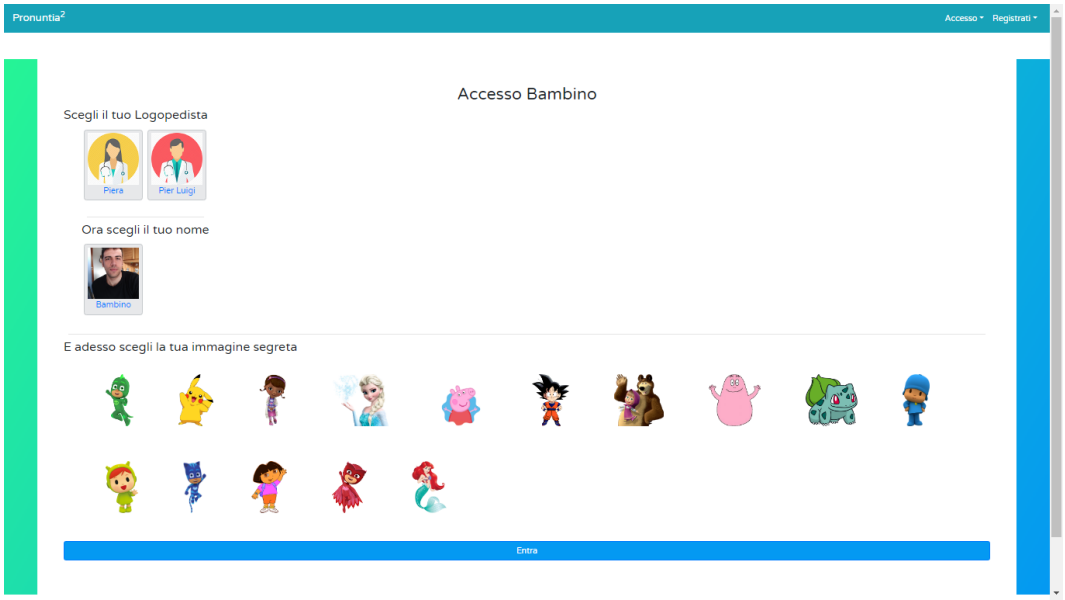

Figure 32 - Patient's login method

Once successfully logged in to the system, it will show the Patient a home screen and play a welcome message, where the

Patient is greeted by name to increase his engagement level with the system.

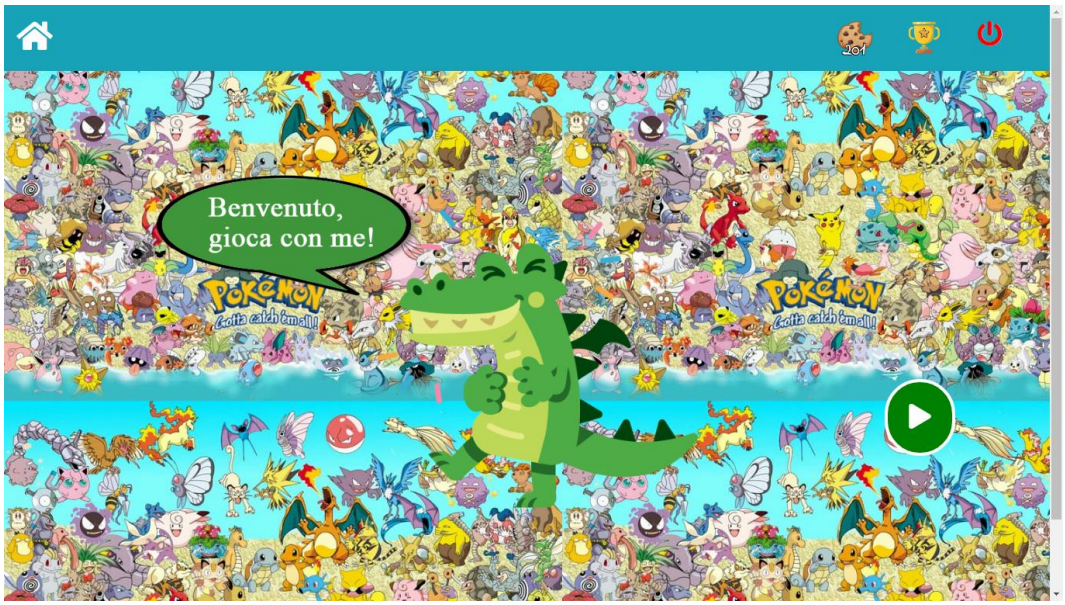

Figure 33 - Patient main page

Since the user who logged in to the system has changed, as it happened for the speech therapist and the caregiver, the navigation bar has changed its appearance for the patient too: here, in fact, it assumes a more playful graphic interface, with colored icons and a different font for the displayed text.

3.2.4.2. Path with exercises

At this point, if there is a therapy in progress in the system for the Patient who has logged in, it will show him/her the exercises of

which the current daily therapy is composed: in this screen, called "path", there will be as many numbered dots as there are the exercises that he/she has to do during the day of therapy, as if they were the levels to be passed in a game path. If the child selects an exercise still to be performed, recognizable by the background of the dot that continuously changes its color, the system shows him a window with a message indicating which exercise he is going to perform together with a presentation of the same, to explain - or remember - to the child what the game is about.

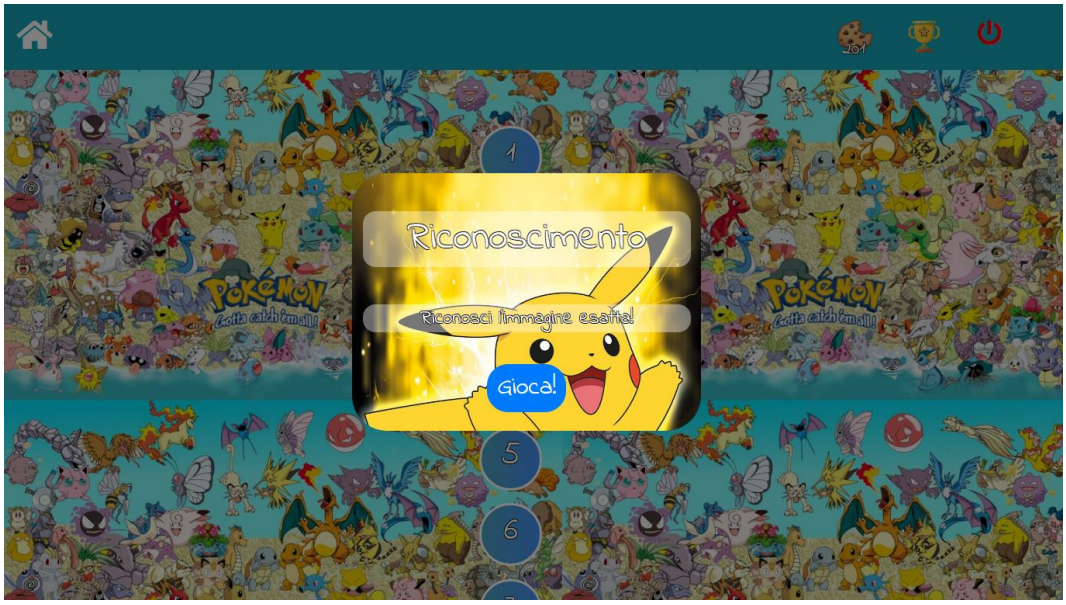

Figure 34 - Game path with explanation message

3.2.4.3. Exercises execution

Depending on the type of exercise the child is doing, the proposed interface changes.

For the first typology of exercise, Naming Images, the system will show on the left the image associated to the word to pronounce, and on the right an orange button with the icon of a microphone: under this button, if inserted by the speech therapist in the creation phase of the exercise, the hint for the execution of the exercise will be present in the form of buttons representing a loudspeaker; pressing on them, the system will reproduce vocally the selected hint.

To perform the exercise, the child must press on the microphone icon, pronounce the word, and when he has finished, press again on the microphone.

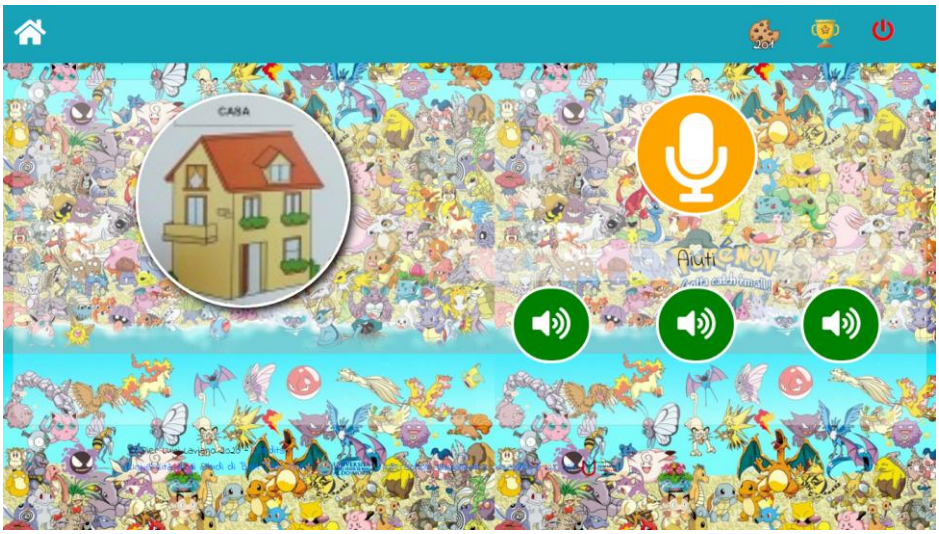

Figure 35 - Execution of Naming Images exercise

If the second type of exercise, Minimum Pair Recognition, is selected, 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, associate it with one of the two images and press on the correct one.

Figure 36 - Execution of Minimum Pair Recognition

For the last type of exercise, Repetition of Words, two buttons will be shown in the center of the screen, the first one representing

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 themselves 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 must press the button with the microphone icon, pronounce the sequence of words in the correct order, and press the same button again.

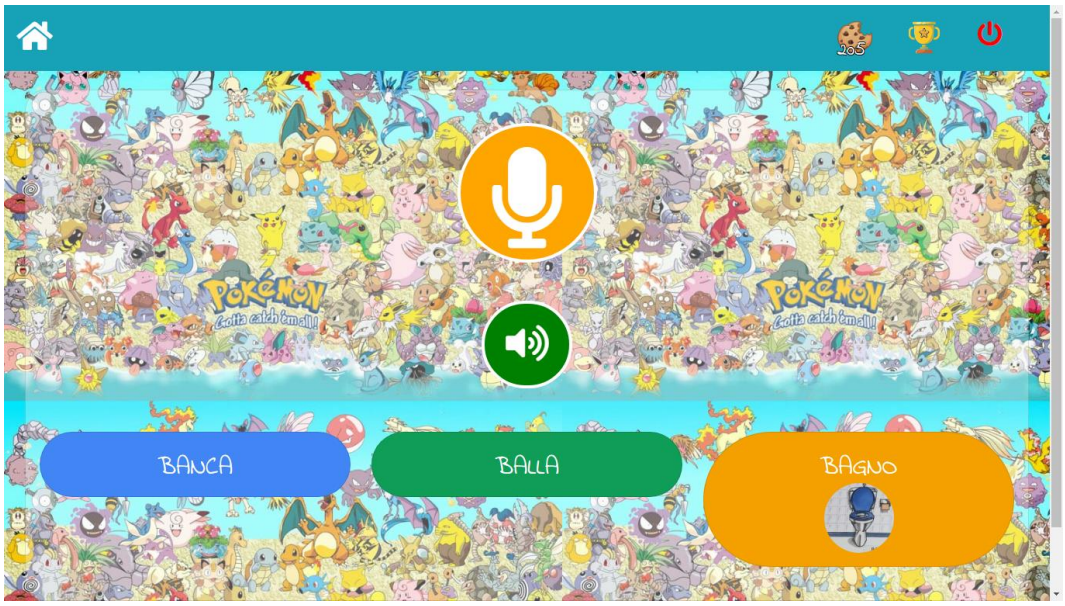

Figure 37 - Execution of Repetition of Words

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 successfully completing the exercise, regardless of the outcome. In case the exercise is performed correctly, the message will also indicate the cookies he earned.

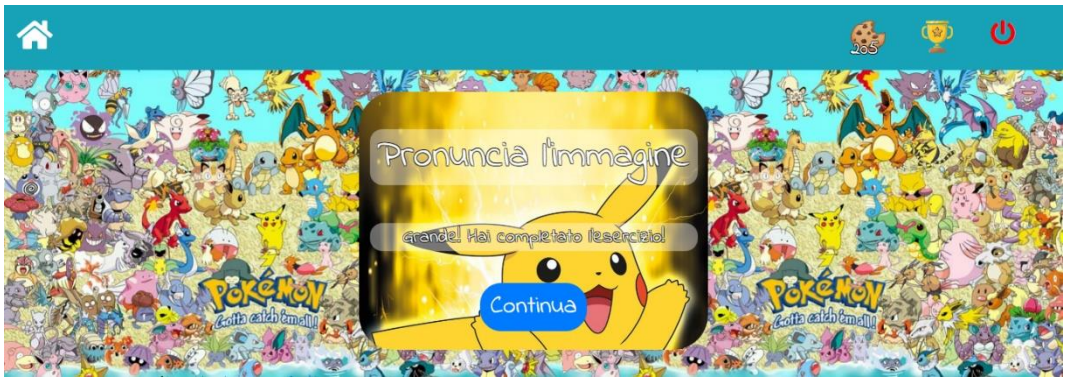

Figure 38 - Exercises executed message

CHAPTER 4

System experimentation

INTRODUCTION

A usability test has been performed to evaluate Pronuntia[2]. It was mainly aimed at evaluating the interaction with non-expert users of the domain.

4.1. Experimentation design

As a sample of users, eleven not expert people in the specific domain of speech therapy with no IT skills have been selected in order to obtain as much heterogeneous feedback as possible on the actual easy usability of the system.

The participants were asked to perform two sets of tasks: the tasks of the first set were about operations to be carried out using the system from the perspective of the Speech Therapist, while the second concerned operations carried out by a Caregiver, associated with a Patient of the former.

Given the health regulations in place due to the spread of COVID-19, it was decided to carry out this phase of experimentation remotely, allowing participants to perform the test from their homes using teleconferencing software for interaction between them and

myself, moderator of the experiment. At the end of the execution of both sets of tasks, it was asked the participants to fill out two questionnaires, submitted online, one for each type of user who they have interpreted, to evaluate individually the experience with the use of different functionalities of Pronuntia[2]. Subsequently, a specific analysis was carried out on the tasks, considering the behavior of the participants and the feedback they gave back during the execution and quantitative analysis of the results obtained from the questionnaires.

4.2. Execution

The experiment was conducted individually with each participant. Each of them was asked to participate in a video call using the Google Meet teleconferencing application, activating the webcam of the device they chose to use and the screen sharing feature, to track, respectively, any facial expressions that could reveal spontaneous feelings and the actions they took to try to complete the assigned tasks.

At the beginning of each test, to each participant was briefly listed the rules for carrying out the experiment; in fact, they were told

that after a short overview of what they would see, they would be asked to use the system spontaneously, letting the system guide them to achieve the goals of different tasks that I would present from time to time. Participants were not told that the execution of each task would be timed to measure their performance, to avoid them feeling under strict examination and, therefore, going quickly in search of the correct solution path, most likely leading to failure and, above all, not understanding what they were accomplishing.

As soon as he reached the system's website, each participant was described Pronuntia[2] as a web system created and designed to support Speech Therapists who have in charge patients to treat remotely, followed in therapy by their Caregivers; this was done to make each participant immediately aware of who are the users who use a system of this type, and from which perspective they should use and evaluate the system.

Initially, each user was asked to register and login to the system at first as a speech therapist and then as a caregiver through the specific sections and, once these preliminary steps were taken, a general overview was made of the features that the specific figure

benefits from, which would be, subsequently, the subject of the tasks to be performed.

Before administering to each user the task to be carried out, defining the objective that was required to achieve, the description of that specific functionality was repeated, to reinforce the understanding of that concept and make him more immersed in the specific context.

During the execution of each task, each participant spontaneously adopted the technique of thinking aloud, allowing, together with the tracking of facial expressions and mouse movements, to be able to take notes to reconstruct more accurately the interaction that took place between the user and the system; subsequently, at the end of each task or subset of tasks, each user was asked for feedback on the arrangement of elements on the screen, whether they considered it appropriate to the task they were performing, and on what they had just performed, asking them whether in their opinion they had achieved the required objective and whether the system had behaved consistently with their intentions and actions.

In addition, after the execution of each task, each participant was given the context of the next one and was asked to try to reach the section independently to perform it, specifying that it would not be evaluated; this, actually, was done in order to get further clues regarding the user's effective understanding of the system, and to be able to evaluate whether the on-screen arrangement of the elements and their indication was sufficiently useful to move around the system without the help of an expert of the same (e.g., while they were in the section showing all the Exercise Series present in the system, they were asked "Now, try to view the detail of the Series you created", or "Now go to the homepage of the system").

Finally, at the end of the completion of all the tasks, the participant was informed that the test was over and was sent two questionnaires to fill out, later and privately, with the maximum sincerity about their experience with the system first from the perspective of the Speech Therapist and then from that of the Caregiver.

4.3. Analysis of the results

Eleven users between the ages of 21 and 26 (average age 24) participated in the system testing phase, including three female and eight male users. Ten participants decided to use a PC with Windows 10 operating system and Google Chrome or Microsoft Edge as web browsers to conduct the experimentation, and only one person a device with macOS Sierra operating system and Safari web browser. Approximately 45% of the sample rated their computer skills very positively, indicating scores of 9 or 10 on an increasing 10-value numerical scale; in each case, only two people responded with an insufficiency (5) to this question.

The participants were administered two sets of tasks, one for the Speech Therapist perspective of the system, and the other one for the Caregiver. They were considered successful if accomplished within 5 minutes. In the *Table 4* and *Table 5* the details of both the sets of tasks and the relative performance are shown.

Tasks for the Speech Therapist		
ID	**Description**	**Average time (minutes)**
ST1.	Creation of a new Word with public visibility	02:01
ST2.	Creation of a new Word with private visibility	01:13
ST3.	Creation of a Naming Images Exercise	01:33
ST4.	Creation of a Minimum Pair Recognition Exercise	01:05
ST5.	Creation of a Repetition of Words Exercise	01:16
ST6.	Creation of a Series of Exercises	01:39
ST7.	Add a Patient to the system	01:39
ST8.	Create and assign a Therapy for a Patient	02:32
ST9.	Explore and set new threshold for your Patients	01:10

Table 4 - Tasks for the Speech Therapist

At the end of tasks ST1. and ST2., two users showed perplexity regarding the system's request for an image to be associated with the word: probably, this was caused by an unclear explanation of the concept of Word itself, which must necessarily be combined with an image in order to use it within the exercises proposed within the system. Despite this, they still managed to complete the tasks successfully, guided by the error messages returned by the system.

For the tasks that concerned the creation of the different types of Exercise, respectively ST3., ST4. and ST5., most of the users found difficulties in using the form useful to search for Words in the

system: in particular, they thought they had to enter the phoneme used in the Word in the search bar but did not understand that they had to press the "Search" button immediately below in order to filter the results that would then appear in the drop-down called "Filtered words list". One user, the one who used the device with macOS and Safari web browser, on the other hand, had correctly guessed the use of the search form, but the system was not working since when she correctly pressed the button to filter the results, the filtering process did not take place. However, these problems did not prevent any user from completing the tasks successfully.

The task related to the section of the Speech Therapists that took the longest time on average for its execution was ST8; the reason for this is to be found in the fact that in this section the participants found the exercises created shortly before too and, getting involved with the system, they wanted to have fun personalizing the therapy as they saw fit.

Tasks for the Caregiver		
ID	**Description**	**Average time (minutes)**
C1.	Complete the registration in the system and login	01:36
C2.	Set a new password for your child	00:24
C3.	Set a new profile image for your child	01:05
C4.	Create and select a new Scenario for your child	02:07
C5.	Logout from the system and login as a Patient	00:55
C6.	Analyze your child's Statistics	00:13

Table 5 - Tasks for the Caregiver

For the task C1. none of the participants found any difficulties, except one of them that did not receive, apparently, the e-mail with the instruction to register into the system: after a brief investigation, he found out that the e-mail provider he used considered the e-mail sent from Pronuntia[2] as a spam message, and the Microsoft Windows 10 Mail application client he used to check his e-mail did not let him to see the spam messages folder. He solved the problem checking the e-mail from the web browser.

In the task C5. they were asked to logout from the system, leaving the Caregiver perspective for a moment, and try to perform the login as if they were the child they take care of: in performing this task, about the half of the participants accomplished it at a first

try, while the others had difficulties because they thought that the first user list that appears, despite the heading that says “Choose your Speech Therapist”, were showing the list of children registered into the system. This could be probably due to a mispositioned instruction label for this section, and certainly a lack of knowledge of the system they were using.

For all the other tasks not mentioned, instead, no particular problems were found in the experimentation phase of the system and, in general, during the performance of the tasks, the users themselves spontaneously returned feedback, above all positive, on the positive experience they were having.

The first section of both questionnaires administered to the participants consisted of the AttrakDiff[10], a tool used to measure the attractiveness of interactive products. Through an evaluation of opposite pairs of adjectives, users (or potential users) of the system can indicate their perception of the product, leading to the evaluation of four dimensions: Pragmatic Quality (PQ), which describes the

[10] *http://attrakdiff.de/*

usability of a product and indicates how successful users are in carrying out correctly their goals while using it; Hedonic Quality - Stimulation (HQ-S), which indicates the extent to which the product can support in the user needs to develop and "move forward" in terms of new, interesting and stimulating functions, content and styles of interaction and presentation; Hedonic Quality - Identity (HQ-I), indicates the extent to which the product allows the user to identify with it; Attractiveness (ATT), which describes an overall value of the product based on perceived quality.

From the first series of questions, those aimed at assessing the perceived actual usefulness of using the product, it emerged that the users who tested the system found the Speech Therapist side slightly more confusing than the Caregiver side, even though the Caregiver side was considered more complicated: as a result also of the analysis of the feedback returned by the participants, the cause could be found in the fact that the Speech Therapist has access to numerous functionalities within the system, which are well structured but can lead a first-time user to confusion; on the contrary, the side of the system dedicated to the Caregiver has fewer functionalities available, but they were perceived as disconnected from each other, which led

them to find it difficult to follow a unique flow of execution. However, they felt that the entire system had a minimalist style, considering that each feature had the essential elements to successfully perform the tasks.

From the subsequent questions, regarding the stimuli that the system conveyed to the users who tested it, it emerges first of all the professional and elegant style perceived by the users: being a web system dedicated to Speech Therapists, therefore health professionals, it was important to obtain a good result under this aspect. Given the large number of features that allow the customization of the system made available to the Speech Therapist, both from a strictly personal point of view and concerning the Patients in their care, the participants in the experimentation have considered the system generally quite engaging (a majority of responses +2); the same cannot be said from the perspective of the Caregiver, who despite having obtained average positive evaluations, was also evaluated negatively (two responses -1), probably due to a good opportunity to personalize the system (the password-image to choose with the child and the creation of the Scenario), but in a

significantly lower number than those offered to the Speech Therapist.

For the last dimension, consistent results were obtained between the two questionnaires administered to participants; in fact, the trend of responses to the proposed questions was almost the same, indicating consistency of identity across the entire system. Participants in the experiment found the system overall useful and the experience with it compelling, considering it attractive and easy to use in continued use.

A later section of the questionnaire consisted of the NASA-TLX [16], that is a test designed to assess the cognitive load of users using a system. The concept of the NASA-TLX considers a task based on recall: in fact, the user is forced to rely on what he remembers of the effort he made in order to perform tasks successfully, so he must rely on his memory of the work he did. Such a test is administered by asking a series of questions about the effort and frustration involved in using the system and the required response are numerical values on a scale of 1 to 10.

From the aforementioned test, it emerged that from the Speech Therapist's perspective, users performed significantly more mental

effort than physical effort to achieve the required goals; from the Caregiver's perspective, on the other hand, almost all users expressed no particular trouble, either mental or physical. This result is in line with the analysis made previously, which show, on the one hand, that the system associated with the figure of the Speech Therapist offers many features that a person not familiar with the system tends to be difficult to fully understand right away, and on the other hand, from the perspective of the Caregiver, the lower number of features available leads to less confusion; in both cases, in any case, the low physical effort required indicates the good organization of the elements in the graphical interface, which allowed users to move easily between the sections of the system.

The next section presented the UeS Short-form questionnaire, designed to measure user engagement while using the system. The questions that comprise it are in the form of a five-value Likert scale, and the user is asked to respond between "disagree" and "agree". From the analysis carried out in this section, it emerged that the degree of involvement perceived by the users can be considered quite high, since almost all the people gave positive answers to the proposed questions, for both questionnaires. In general, all users

considered the experience of using Pronuntia[2] very rewarding and interesting, consistent with the results that emerged from the previous sections.

Conclusions and future works

Pronuntia[2] is a web 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 him in the execution of therapies.

This work reports the elicitation and definition of new requirements obtained starting from a first prototype of Pronuntia, deeply re-designed in all its aspects. Automatic correction of exercises has been improved by introducing a mechanism of thresholds customizable by speech therapists, who can make the system more or less sensitive for correction for each patient category. In addition, machine learning techniques were analysed and an automatic classification application was implemented to allow the system to perform a screening of new (and pre-existing) patients to determine the severity of their level of language impairment. A login method for children has been implemented, to make them autonomously access the system in an easy and secure way.

From the experimentation phase of the system carried out with eleven participants not experts in the domain, resulted that the system allows users to carry out adequately and without excessive effort all

the tasks required by who uses it with a high degree of efficiency, and arouses gratification and interest in those who used it.

The system, as future works, has some aspects that can be improved to reduce the confusion perceived in various sections of the system, with the introduction, for example, of more meaningful labels to indicate the different functions of the system, or an improvement of the authentication technique used by patients, which, although functional, is not very convincing in terms of privacy.

Bibliography

[1] S. Crosbie, A. Holm, and B. Dodd, “Intervention for children with severe speech disorder: a comparison of two approaches.,” *Int. J. Lang. Commun. Disord.*, vol. 40, no. 4, pp. 467–491, 2005, doi: 10.1080/13682820500126049.

[2] American Speech-Language-Hearing Association, “Definitions of communication disorders and variations,” 1993.

[3] F. G. Gall, “Works: On the functions of the brain and each of its parts.” 1835.

[4] L. B. Leonard, “Children with specific language impairment and their contribution to the study of language development.,” *J. Child Lang.*, vol. 41 Suppl 1, no. 0 1, pp. 38–47, Jul. 2014, doi: 10.1017/S0305000914000130.

[5] A. L. Benton, “Developmental aphasia and brain damage,” *Cortex*, vol. 1, no. 1, pp. 40–52, 1964.

[6] “Aphasic children. Identification and education by the association method. By Mildred A. McGinnis, Director emeritus, division of speech correction and pathology, central institute for the deaf, St. Louis, Missouri. Foreword By S. Richard Silverman, Director, central institute for the Deaf. Alexander Graham Bell association for the Deaf, Inc., Washington, D. C., 1963, pp. 154, 20 illustrations,” *Laryngoscope*, vol. 74, no. 3, p. 433, 1964, doi: https://doi.org/10.1288/00005537-196403000-00009.

[7] H. R. Myklebust, “Language Disorders in Children,” *Except. Child.*, vol. 22, no. 4, pp. 163–166, 1956, doi: 10.1177/001440295602200406.

[8] M. E. Morley, *Cleft Palate and Speech*. E. & S. Livingstone, 1954.

[9] N. Chomsky, *Syntactic structures*. The Hague: Mouton, 1957.

[10] B. Wulfeck, E. Bates, M. Krupa-Kwiatkowski, and D. Saltzman, “Grammaticality sensitivity in children with early focal brain injury and children with specific language impairment,” *Brain Lang.*, vol. 88, no. 2, pp. 215–228, 2004, doi: 10.1016/S0093-934X(03)00100-7.

[11] A.-M. Oster, “Quarterly Progress and Status Report Presentation of a new EU project for speech therapy : OLP (Ortho-Logo-Paedia),” *Proc. Fonetik, TMH-QPSR*, no. June 2002, pp. 45–48, 2002.

[12] K. Vicsi, P. Roach, A.-M. Öster, Z. Kacic, P. Barczikay, and I. Sinka, *SPECO - a multimedia multilingual teaching and training system for speech handicapped children.* 1999.

[13] J. P. Haton *et al.*, “Speech Training for Deaf and Hearing-Impaired People,” no. December 2013, 1999.

[14] D. Beukelman and P. Mirenda, *Augmentative and Alternative Communication*. 2012.

[15] L. M. T. Jesus, J. Martinez, J. Santos, A. Hall, and V. Joffe, “Comparing Traditional and Tablet-Based Intervention for Children With Speech Sound Disorders: A Randomized Controlled Trial,” *J. Speech, Lang. Hear. Res.*, vol. 62, no. 11, pp. 4045–4061, 2019, doi: 10.1044/2019_JSLHR-S-18-0301.

[16] J. P. Tracy and M. J. Albers, “Measuring cognitive load to test the usability of web sites,” *STC’s 53rd Annu. Conf. Proc. 2005*, no. October, pp. 256–260, 2006.

RINGRAZIAMENTI

Terminato questo lavoro di tesi, ci tengo a ringraziare le persone che mi sono state vicine e l'hanno reso possibile.

Innanzitutto, un ringraziamento speciale va al mio relatore, il professor Antonio Piccinno, per avermi guidato durante questo percorso e per avermi trasmesso l'impegno e la passione da mettere in tutto ciò che si fa.

Ringrazio il professor Stefano Ferilli per aver supervisionato il lavoro svolto per questa tesi riguardante la Sua materia.

Un sincero ringraziamento a Isabella Roppa e Roberta Cassano per avermi fatto addentrare nel mondo della Logopedia e per il generoso impegno che hanno profuso nello sviluppo di questo progetto.

Ringrazio i miei genitori e le mie sorelle Ornella e Morena per esserci sempre stati per me in qualsiasi momento e, anche con una sola parola, sono sempre stati in grado di farmi spronarmi.

Un enorme ringraziamento va ad Alberto, il Rosso, Dota, Peppe, Raf, Bobby, Umbertone e Valentina per la spensieratezza che mi infondono quotidianamente con le loro personalità uniche.

L’ultimo ringraziamento, probabilmente quello più sentito, è per Alessia, senza la quale questa tesi non sarebbe mai esistita perché non avrei mai avuto il coraggio di continuare il mio percorso universitario. Grazie per esserci sempre per me, sempre lì in prima fila a spronarmi a dare il massimo e per essere un bellissimo esempio da seguire. Grazie per l’infinita serenità che riesci a trasmettermi e con cui viviamo il nostro rapporto, da sempre ostacolato dalla distanza, ma che mai nulla riuscirà a dividere. Sei da sempre il mio specchio, in cui guardo per cercare la parte migliore di me che soltanto tu riesci a tirare fuori.