Ramses: A Robotic Assistant and A Mobile Support Environment for Speech and Language Therapy

Vladimir Robles-Bykbaev,Tania Flores-Tapia, Celia Ordóñez-Arce
GIIATa, Universidad Politécnica Salesiana
Cuenca, Ecuador
Martín López-Nores, Jorge García-Duque
José Pazos-Arias. Spain
vrobles@ups.edu.ec, mlnores@det.uvigo.es, jose@det.uvigo.es
jgd@det.uvigo.es, tflores@est.ups.edu.ec, cordonezar@est.ups.edu.ec

ABSTRACT: *Speech and Language Therapy (SLT) is an area focused in the rehabilitation of people suffering from different kinds of disorders and disabilities related with language and communication. According to latest estimates of the World Health Organization, the most of countries do not have appropriate structures, and enough personnel and resources to provide healthcare and rehabilitation services these persons. This problem is more complex in developing countries, given that a Speech-Language Pathologist (SLP) must to carry out several activities besides providing the therapy as such. On those grounds, this paper presents a robotic assistant able to provide support for SLPs during the execution of several activities related with the SLT such as: patient's progress monitoring, interacting with patients suffering from emotional and behavioural disorders, report generation, and in general, supporting the execution of different kinds of therapy activities (visual, motor, and auditory stimulation, relaxation activities, etc.). This approach is based on an integrative environment that relies on mobile ICT tools, an expert system, a knowledge layer and standardized vocabularies. This proposal has been tested on 65 children suffering from different types of disabilities, and the achieved results are encouraging.*

Keywords: Expert System, Mobile Applications, Robotic Assistant, Speech-Language Therapy

Received: 12 May 2015, Revised 11 June 2015, Accepted 18 June 2015

1. Introduction

The language and communication abilities constitute one of the cornerstones of cognitive development of children, since they provide a set of skills that any individual can use to interact with his/her environment, or express his/her needs, ideas, thoughts, and feelings. However, currently an important number of persons in the world (children and adults) suffer from different kind of disabilities that can affect their language and communication skills. Some of the latest world estimates present a complex overview, with 15% of the world's population living with some form of disability, whereas 60 million of persons live with disabling hearing loss and 15 million suffer from stutter. In the same line, the number of existing SLP confirms a lack of personnel to provide adequate healthcare and rehabilitation services: in sub-Saharan Africa there is 1 SLP for every 2-4 million people, while in the US, UK, Austria, and Canada there is 1 SLP for every 2500-4700 people [7].

Commonly, People With Communication Disorders (PWCD) can be affected by a sability. The range of the existent disabilities is large, and a disability can appear in any stage of the person's life. In the early life stages, a delay in meeting evelopmental milestones may be secondary to perinatal events, involving complicated interactions between mother and fetus during delivery. Maternal factors including weight, diet, and morbidities can affect neonatal adaptation and later development. Prematurity, low birth weight, and previous intrauterine insults as well as complications. during delivery of a previously normal fetus increase the risk for perinatal stress [5]. Some of the most common speech and language disorders are the following [1]:

• **Speech disorder** is present when a person is not able to produce speech sounds correctly or fluently, or has problems with his/her voice.

• **Language disorder** appears when a person cannot understand in a proper way other persons (receptive language) or sharing thoughts, ideas or feelings (expressive language).

• **Swallowing disorder** can occur during some of the three phases of swallowing process (oral, pharyngeal or esophageal).

• **Hearing disorder** appears when damage exists in the auditory system. This disorder can be one of these three types: conductive hearing loss, sensorineural hearing loss, or mixed hearing loss.

These disorders can be related or not with several types of disabilities/diseases like Cerebral Palsy, Alzheimer's disease, Parkinson's disease, bacterial meningitis, intellectual disabilities, Autism Spectrum Disorder, etc. [1]. On those grounds, in this paper we present a comprehensive approach to support the different activities conducted by SLP, with the aim to provide a better healthcare and rehabilitation services for PWCD. Our approach uses an integrative ICT layer of robotic, mobile and web tools that relies on formal knowledge model (archetypes and ontologies) and standardized vocabularies. In the next sections we will describe in detail the most important features of the Robotic Assistant and the Mobile Support Environment for SLT (RAMSES).

This paper is organized as follows. In Section 2 are presented some relevant researches related to ICT tools and SLT. A general overview of the proposed approach and the details related with the robotic assistant and the mobile environment are described in Section 3. The pilot experiments conducted in order to validate and analyze the patient's response to RAMSES are shown in Section 4. Finally, Section 5 presents conclusions and some lines of future work.

2. RelatedWork

In the last decade several researches have been conducted with the aim of improving different activities conducted during the SLT. In the rehabilitation area of patients with stroke, [4] proposes the use of humanoid robot to provide speech and physical therapies for a patient suffering from aphasia and hemiparesis. In this research were analyzed two treatment strategies: the "sole condition", and the "sequential condition". The first one consists on applying speech therapy during the first 4 weeks, and physical therapy during the 6 next weeks. In the second strategy is applied a combination of speech and physical therapy during 6 weeks followed by a 4-week post-treatment period. Based on the results of this study it is possible to analyze the fatigue effect due to the combination of physical activities and speech therapy, the improvements achieved during the speech therapy, the adverse effects of upper-limb exercises on subsequent naming tasks, as well as other many variables [4].

On the other hand, several older adults suffering from dementia have shown recovery from depression, reduction of agitation, and recovery from speech disorders during therapies conducted with support of PARO robot (a seal pup) [8]. The main idea behind this proposal is based on previous researches done in the field of psychology, that have proven the great benefits of working with animales during therapy sessions [2]. In general, through the "animal therapy" is possible to achieve three effects: psychological effect (relaxation, motivation, concentration, etc.), physiological effect (improvement of vital signs as cardiac rhythm, breathing, etc.), and social effect (improvement of predisposition to start communication processes with caregivers, other patients, etc.). In this line, PARO is a robot able to recognize words, senses where and how the patient touches its body, and responds accordingly. This robot was tested in several clinics and rehabilitation centers of Japan, the European Union, and United States since 2005. The results of using PARO have shown important improvements in behavioural problems and depression of elderly suffering from dementia [8].

The interaction between children with special needs and robotic interviewers is studied in [11]. The aim of this research was to

analyze the children behavior during the interviews conducted by robot compared to interviews done by human interviewer. The results reveal very similar outcomes in the children responses for both cases (amount of words, filler words and key words). Other approaches propose the use of robots to support therapy sessions and improve the social interaction skill in children with Autism Spectrum Disorders (ASD) [10, 9], motivate children with communication disorders to learn the sign language [6] or provide support for Mexican patients suffering from motor speech disorders (dysarthria) [3].

The aforementioned proposals do not pay attention to the following issues of management of healthcare and rehabilitation: the modeling of domain knowledge and the use of international standards and data structures to store, manage, and share the patients' health records.

3. The Proposed Approach

The proposed model relies on several services, knowledge structures, and functionalities that are depicted in figure 1. Some of the most relevant for this paper are the following:

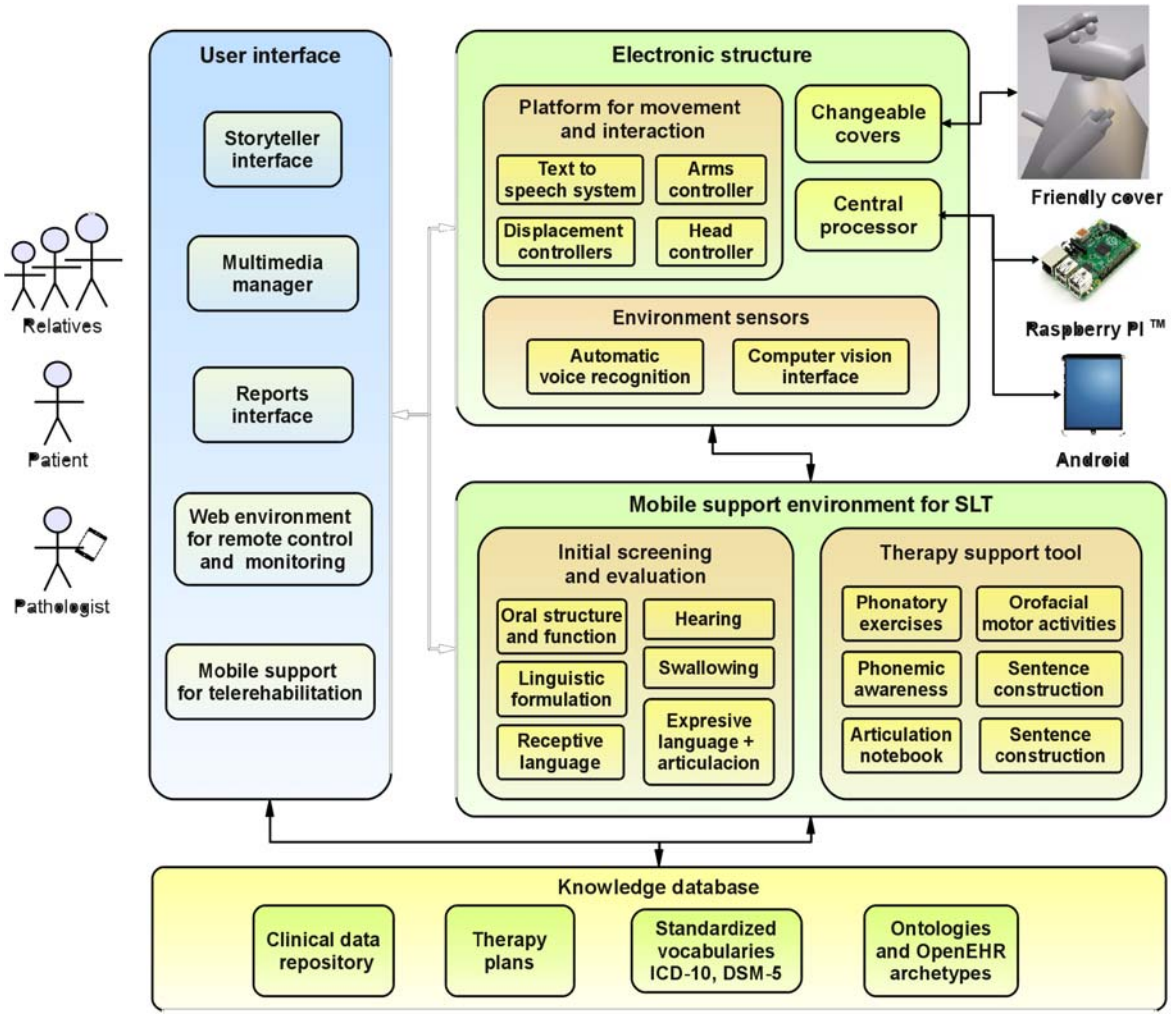

Figure 1. The general components tha constitute the proposed system.

• **Actors**: our proposal considers all people involved in the SLT process. The model helps doctors and SLPs with different kind of activities related with therapy and health care like report generation, support during therapy sessions, patients monitoring, and multisensory stimulation. The relatives can use several mobile applications to conduct reinforcement activities in home and help with patient's monitoring, whereas students (future SLPs) can be trained using real cases and automatized tests.

• **User Interface** with the aim of having a comprehensive, and fully-interactive approach, our proposal implements a layer that has several tools and interfaces to make easiest performing different kinds of operations, according o each user's profile.Some of the most relevant functionalities have been developed to provide support during different tasks as: reports generation, managing of all multimedia resources of the system, remote controlling the robot (via a web interface), storytelling using the robot, and supporting several activities for reinforce the therapy (at home).

• **The Electronic Structure** has been designed bearing in mind several factors as flexibility, easily maintenance, low cost, and possibility to quickly incorporate new functionalities. Therefore, the robot was developed using the same electronic structure for displacement but with two different approaches of his central unit control. In the first case, we have used an Android program to implement the control and interaction system, and this program can be easily installed in any Android compatible device (tablets or smartphones). For the second model, we have used an embedded electronic system based on Raspberry PI and Raspbian, aspect that allows us to reduce significatively the robot cost (from 800 to 530 USD, approximately). Some of the details of this layer, as well as the storyteller will be provided in the section 3.1.

• **Knowledge Database** covers seven main areas of speech and language as hearing, voice, swallowing, receptive language, expressive language and articulation, oral structure and function, and linguistic formulation. Each of these areas allows defining the skills, development milestones, knowledge, assessments and tests that must be considered during the rehabilitation process of PWCD. Through the use of OpenEHR archetypes[1], ontologies, and standardized vocabularies is possible to represent, manage and share all information related with patients/ subjects.

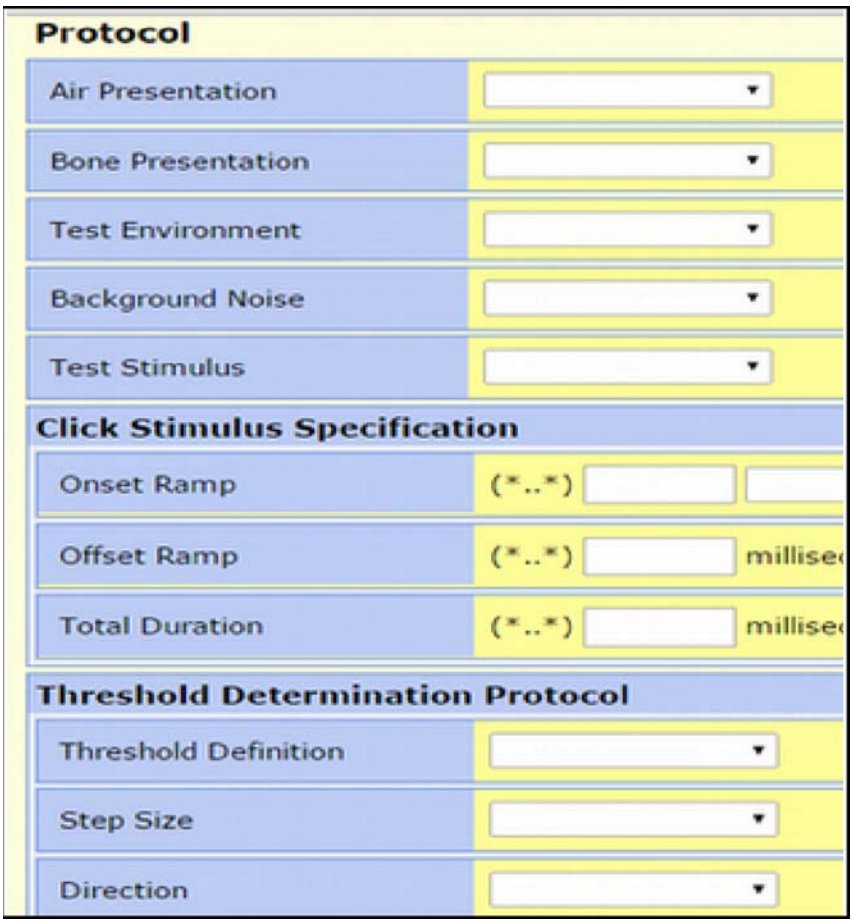

Figure 2. Partial view of the protocol (template) required to conduct a Pure Tone Audiometry Test (developed with OpenGenEHR: https://code.google.com/p/open-ehr-gen-framework).

[1]http://www.openehr.org

Figure 2 shows a partial screenshot of the protocol (template) to conduct a Pure Tone Audiometry test (PTA) test. As we can see, this template is based on an archetype that allows assessing the patient's response to sound stimulation and voice commands, and determine if he/she is able to localize sound sources without visual stimuli. Likewise, Fig. 3 shows a partial view of the archetype that captures the information related with speech fluency patient's evaluation.

Mobile Support Environment for SLT this is a layer where several tools are integrated and the main element is constituted by RAMSES. Some other elements of this layer are the mobile applications to perform therapy exercises, assessment tasks, collection of information related with patients, registering the results of therapy sessions, report generation and sharing data

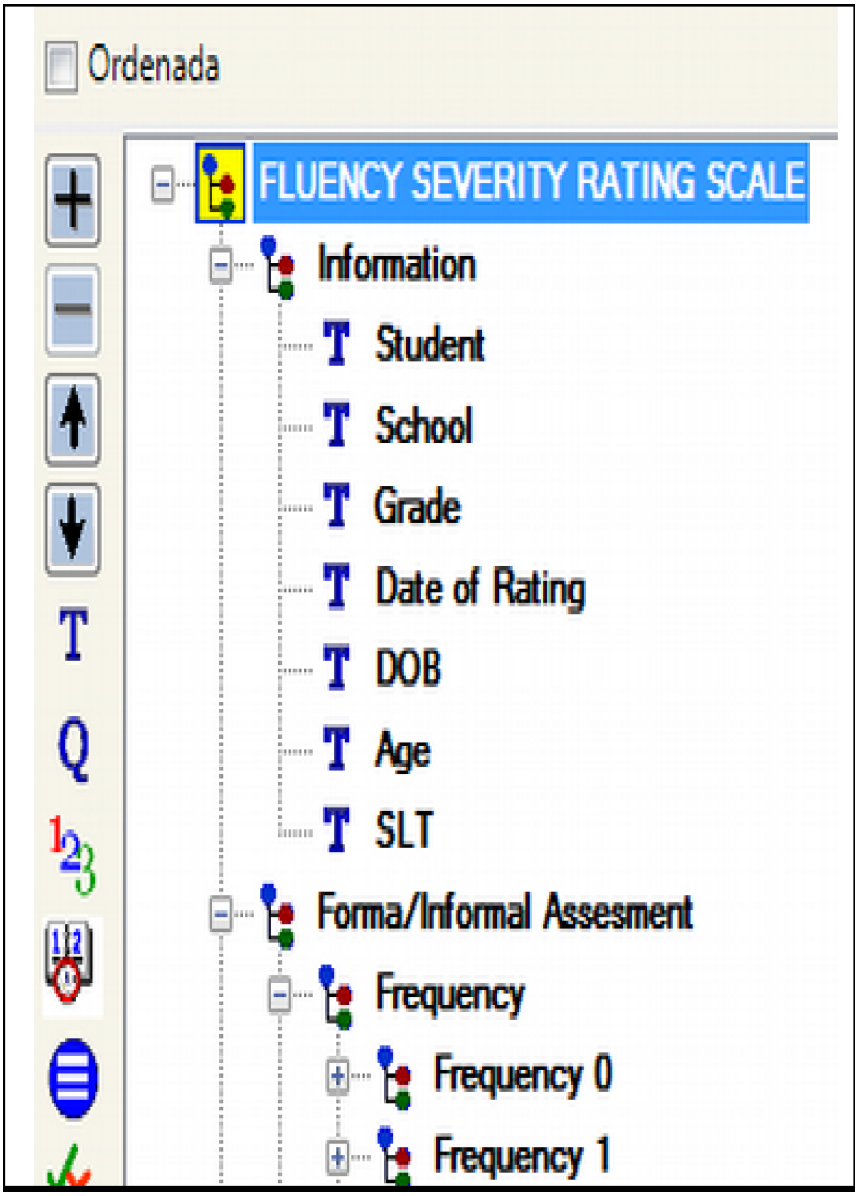

Figure 3. Partial view of the archetype modeled to store the information of fluency evaluation on patients (developed with OpenEHR Archetype Editor).

with other SLP. For example, Figure 4 shows two screen captures of the developed applications. The first one (upper side) is of the mobile application to support the learning process of phonetic code. This tool allows to conduct several kinds of exercises related with articulation (phonemic awareness, sentence construction, phonatory,...). On the other hand, the second one (bottom side) is from the application used to conduct the initial screening, and evaluation of the seven speech-language areas mentioned above.

3.1 Ramses V2

RAMSES is an intelligent environment that uses mobile devices and embedded ectronic systems to provide support during the SLT process. Basically, RAMSES relies in a set of mobile applications, electronic devices, and on a robotic assistant consisting in two elements: a central processor (an Android smartphone or tablet, or an embedded electronic system) and a displacement electronic platform. The robot can perform several activities and provide the following services to SLP:

• Perform movements using his 5 degrees of freedom (displacement on the ground, head rotation, and arms movement in one axis).

• Provide remote control to interact with patients suffering from behavioural disorders as autistic disorder, attentiondeficit, hyperactivity, or other related disorders. In these cases, a SLP can work with children in a Gesell camera and remotely control the robot movements, send text messages that will be readed by the TTS system of the robot, or register the responses provided by patients.

• Reproduce songs, tell stories, search word or sentences in internet using voice commands.

• Recognize four different hand gestures (fist, open hand, closed hand, and semi-closed hand)

• Provide exercises to perform visual, intellectual and motor stimulation on patients (visual discrimination, painting,

• spatial and temporal relationships, auditory memory, ...).

• Detect and recognize objects (of certain colors), and faces from several users.

• Provide a set changeable covers with different forms as dogs, bears, cats, etc., that can be easily developed using a 3D printer.

• The control software can be installed in a wide range of Android OS versions (starting from version 3.0).

• Provide an autonomous functioning of 8 hours with an independent rechargeable battery.

In order to recognize the hand gestures, the patient must to wear a glove of a uniform color (typically, blue, yellow, or green). Once the image is captured, we extract an descriptor that relies on the shape's signature of the polygonal approximation and convex hull. With this combination we can carry out the pattern recognition process with an accuracy over the 90%. Likewise, the knowledge database stores several samples of each gesture, and currently we are using the K Nearest Neighbor criterion to classify the appropriate gesture. The algorithms designed to perform the computer vision tasks were developed using the Open Source Computer Vision Library (OpenCV)[2], aspect that allows us to port the code from the embedded system to mobile device.

3.2 Storyteller

The storyteller is a very useful tool (currently under development) to interact with PCWD, especially with children. This tool is constituted by 6 modules that provide specific functionalities during the short story construction (see Figure 5). Below we describe the most important features of each module.

• In order to interact with the user, the storyteller provides a graphics user interface that contains a touch panel, a user authentication service (with the aim of register the activity performed by patient), an avatar (a animated character that can be a dog, cat, person, tree or any other element), an Automatic Speech Recognition system (ASR) to recognize orders, and a TTS system.

• With the aim of constructing the short stories, the tool incorporates a module with several services of Natural Language Processing (NLP). With these services, the system is able to automatically search short stories, and fairy tales over the internet,analyze the complexity of the texts, extract the constituents of each sentence, and classify each story according to different parameters (complexity, topic, length, etc.).

• Once the robot presents the story to the patient, the tool uses the module of evaluation and monitoring to automatically show exercises related with the story. These exercises can include questions related with some characters or elements of the story, pairing a pictogram with the corresponding word, completing sentences, etc. After the patient has completed the exercises, the tool can perform an automatic evaluation and the corresponding report generation.

[2] http://opencv.org

• With the objective of starting the exercises, the tool provides the storytelling module, which incorporates functionalities to perform the selection of the short story according to patient's profile. The selection can be performed automatically or manual, and depends of the pathologist's decision. Likewise, this module implements a TTS system, that can be configured according to user's preference (voice characteristics as pitch, resonance, prosody, etc.)

• The **central administration module** is responsible of the interaction with/between the other modules, and performs several activities as: response to the touch events (analyze which actions have been done by the user), services to manage information related with users (patients, SLPs, etc.), processing voice commands, and providing a social behaviour (animations for the avatars, response to user requirements, etc.). The **educational contents database** stores all the information related with the users, exercises, reports, and the stories.

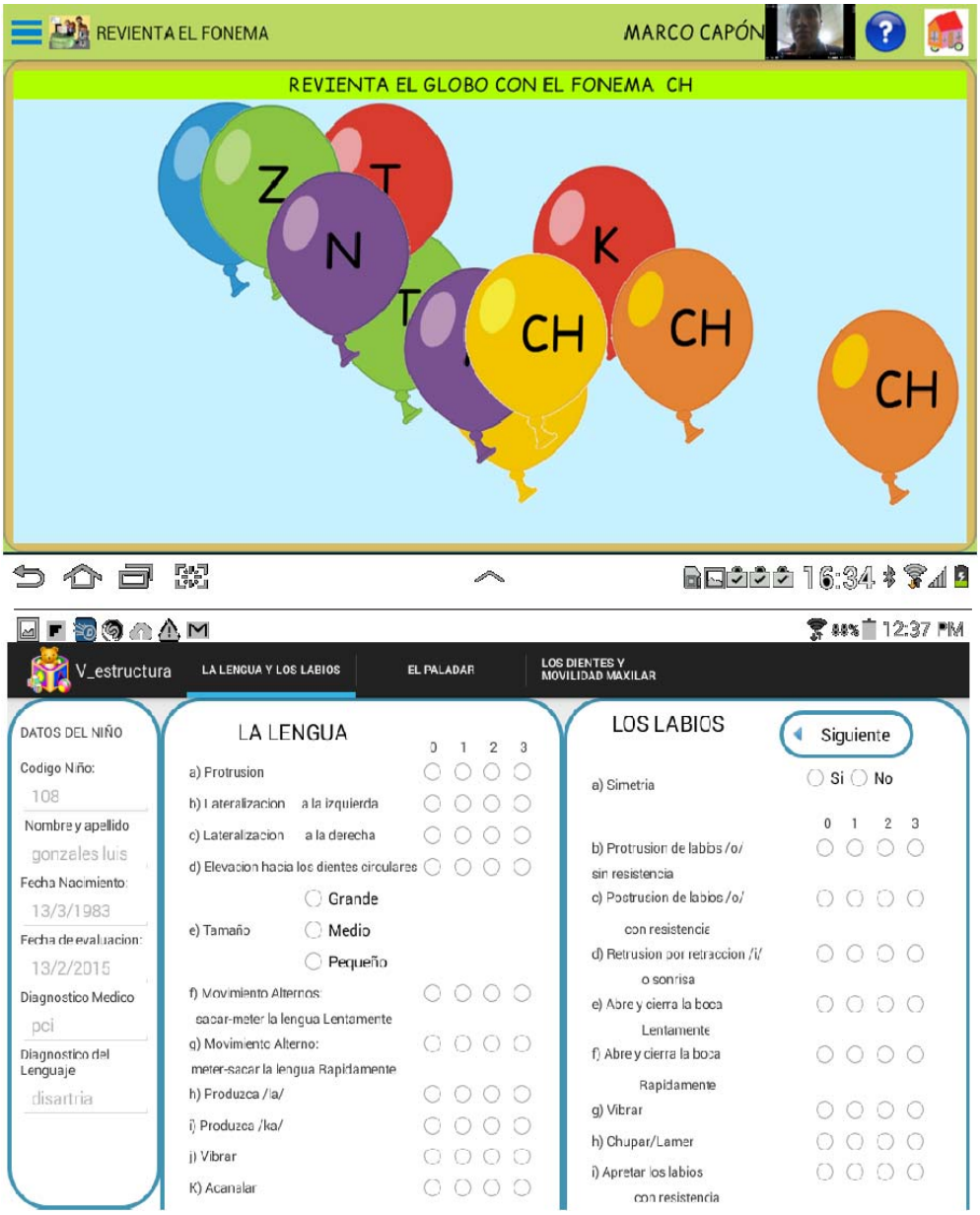

Figure 4. Two examples of the mobile applications developed to provide support during the patient's evaluation, and therapy support. The top screen capture is of the an educational activity of the phonetic learning tool. This consists on the game "pop the balloons" of the given phoneme. The bottom screen capture is of the application to conduct the evaluation on the patient's oral peripheral mechanism

• The **searching module** is an interface that is used by SLPs to search and analyze the stories of the system. With this module, a pathologist can perform an automatic searching according to the parameters mentioned above.

Disability	Number of patients	Average level of interest
Mild Intellectual Disabilities (F70)	23	4.83
Moderate Intellectual Disabilities (F71)	8	4
Cerebral Palsy (G80)	5	4
Down syndrome (Q90)	11	3.55
Autistic disorder (F84.0)	4	3.75
Epilepsy and recurrent seizures (G40.8)	3	4
and Unspecified intellectual disabilities (F79)		
Other disabilities	11	–

Table 1. List of the disabilities, number of patients and average levels of interest achieved by the robot in the different subpopulations. Each disability has the code of the ICD-10-CM standardized vocabulary.

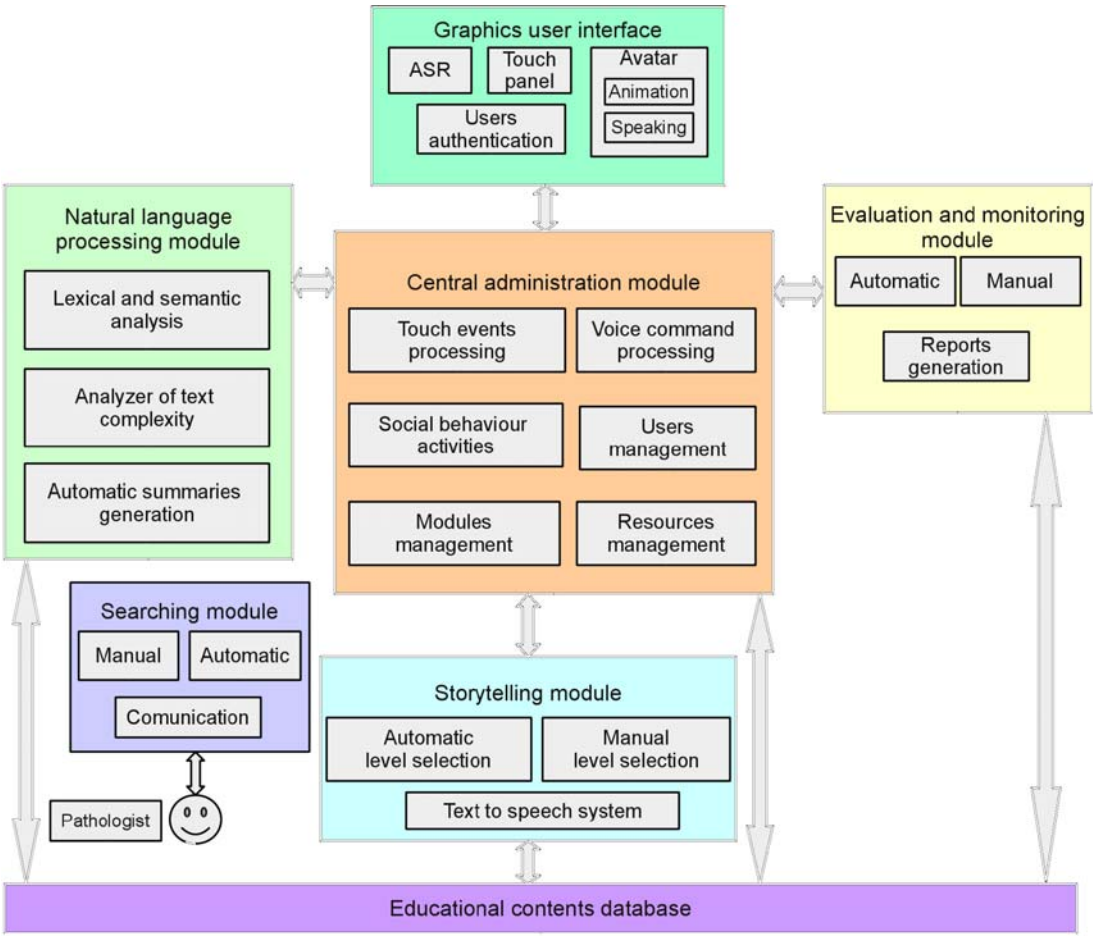

Figure 5.A general overview of the most important modules that constitute the storyteller tool. With the support of this tool, the SLP is able to quickly construct short stories to perform different kinds of activities with patients (especially children).

4. Pilot Experiment and Results

In order to analyze the response of children to robotic assistant, we have conducted a pilot experiment in the Special Education Unit of Azuay - Ecuador (Unidad Educativa Especial del Azuay). During this experiment, the SLPs performed several activities of relaxation previous to therapy, using the applications provided by the robot

• **Auditory Memory Stimulation:** This exercise consists on telling a short story through the TTS system of the robot. Once the robot ends the story, the SLP asks some simple questions related with main characters, events, or objects of the story. According to the responses provided by children is possible to analyze whether they have been aware to storytelling and have understood the main message of the story. Another activity that is carried out is the sounds recognition. The robot plays some sounds of different domestic animals (cow, dog, sheep, etc.) and then the SLP asks the children to point what animal produces that sound.
• **Comprehensive Reading**: Consists of showing some short sentences and then make questions related to the subject of the sentence or the action that takes place within it. According to responses provided by children, the SLPs analyzes if they understand the sentence and the question done.
• **Visual Stimulation**: This activity is based on execution of simple orders as jumping, raise the hands, go to another site, etc., when a certain color appears in the display of the robot.
• **Stimulation of Motor Skills**: The robot presents a virtual painting tool, that responds to some hands gestures. The children can make drawing using the movement of their hands. Likewise, the children must to move basic geometric figures from one place to another of the display.

A total of 65 children suffering from different types of disabilities (Table 1) were evaluated by a team of experts. To determine the level of interest in the robot, the SLPs have used a Likert scale consisting of five degrees (very low = 1, low = 2, medium = 3, high = 4, very high = 5) and analyzed children's response to orders, activities and requirements made by robot. The last column of Table 1 shows the average values of level of interest achieved by the robot in the different subpopulations (the value for 'other disabilities' is not shown, because none was a representative sample).

5. Conclusions

The results obtained with this research show that is possible to automate several activities related with SLT, with the aim to provide a better service to patients suffering from several kinds of disabilities. In the same way, mobile applications constitute important backing tools, allowing SLPs to perform their activities anywhere and in a more comfortable way. Another important point is that throughout the mobile applications the patient's relatives can contribute in a more participative way during the different stages of the rehabilitation processes.

The cost to produce the robot is low, and it allows to conduct a massive production using for it simple materials as a 3d low cost printer, replaceable embedded systems (Raspberry computers, and Arduino electronic cards), transparent plastic, and common electronic elements (step motors, leds, speakers, etc.).

As lines of future work, we propose the following ones:

• Develop a inference mechanism able to automatically select activities or exercises according to the changes in the concentration levels of each patient.

• Design more specific activities and routines according to stimulation requirements for those skills that are affected in each patient.

Acknowledgment

The authors from the Universidad Politécnica Salesiana have been supported by the "Sistemas Inteligentes de Soporte a la Educación, SINSAVE v3" research project. The authors from the University of Vigo have been supported by the European Regional Development Fund (ERDF) and the Galician Regional Government under agreement for funding the Atlantic Research Center for Information and Communication Technologies (AtlantTIC), as well as by the Ministerio de Educación y Ciencia

(Gobierno de España) research project TIN2013-42774-R (partly financed with FEDER
funds). We would like to thank the support provided by the Unidad Educativa Especial del Azuay (UNEDA), and Luis González-Delgado.

References

[1] American Speech-Language-Hearing Association. (2014). International Classification of Diseases, 10th revision - clinical modification, related to speech, language, and swallowing disorders. Technical report, Rockville (MD), USA.

[2] Baun, Mara M., Bergstrom, Nancy., Nancy F Langston, Thoma, Linda.(1984). Physiological effects of human/ companion animal bonding. *Nursing Research*, 33(3) 126–129.

[3]Caballero-Morales, S. O., Bonilla-Enríquez, G., Trujillo-Romero.F. (2013). Speech-based human and service robot interaction: An application for Mexican dysarthric people. *International Journal of Advanced Robotic Systems*.

[4]Choe,Y. K., Jung, H. T., Baird, J.,Grupen, R. A. (2013). Multidisciplinary stroke rehabilitation delivered by a humanoid robot: Interaction between speech and physical therapies. *Aphasiology*, 27 (3) 252–270.

[5]Ergaz, Z., Ornoy, A. (2011). Perinatal and early postnatal factors underlying developmental delay and disabilities. *Developmental Disabilities Research Reviews*, 17 (2) 59–70.

[6] Kose, H., Akalin, H., Uluer. P. (2014). Socially interactive robotic platforms as sign language tutors. *International Journal of Humanoid Robotics*, 11 (1).

[7] McAllister, Lindy., Wylie, Karen., Davidson, Bronwyn., Marshall, . (2013). The world report on disability: an impetus to reconceptualize services for people with communication disability. *International Journal of Speechlanguage Pathology*, 15 (1) 118–126.

[8] Shibata,T. (2012). Therapeutic seal robot as biofeedback medical device: Qualitative and quantitative evaluations of robot therapy in dementia care. *In*: Proceedings of the IEEE, 100 (8) 2527–2538.

[9] Vanderborght, B., Simut, R., Saldien, J., Pop, C.A., Rusu, A.S., Pintea, S., Lefeber, D., David, D.O (2012). Using the social robot Probo as a social story telling agent for children with ASD. Interaction Studies, 13 (3), 348–372.

[10] Villano, M., Crowell, C.R., Wier, K., Tang, K., Thomas, B., Shea, N., Schmitt, L.M., Diehl, J.J.(2011). DOMER: A Wizard of Oz interface for using interactive robots to scaffold social skills for children with Autism Spectrum Disorders. *In:* Proceedings of 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 279–280, Lausanne, Switzerland, March.

[11] Wood, L. J., Dautenhahn, K., Lehmannand Robins, B., Rainer, A., Syrdal, D.S. (2013). Robot-mediated interviews: Do robots possess advantages over human interviewers when talking to children with special needs? Lecture Notes in Computer Science, 8239. 54–63.