Expert Systems with Applications 42 (2015) 7641–7651

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

# SPELTA: An expert system to generate therapy plans for speech and language disorders

Vladimir E. Robles-Bykbaev [a,b], Martín López-Nores [b,*], José J. Pazos-Arias [b], Daysi Arévalo-Lucero [a]

[a] GI-IATa, Centro de Investigación, Desarrollo e Innovación en Ingeniería (CIDII), Universidad Politécnica Salesiana, Calle Vieja 12-30, Cuenca, Ecuador
[b] AtlantTIC Research Center for Information and Communication Technologies, Department of Telematics Engineering, University of Vigo, EE Telecomunicación, Campus Universitario s/n, 36310 Vigo, Spain

ARTICLE INFO

*Article history:*
Available online 14 June 2015

*Keywords:*
Speech–language therapy
Generation of therapy plans
Clustering

ABSTRACT

Speech and Language Pathologists have to treat a wide spectrum of disorders, working with a large set of exercises and activities to design personalized therapy plans for their patients. This paper presents an expert system designed to provide support in that labor by automatically generating therapy plans containing semiannual activities in the areas of hearing, oral structure and function, linguistic formulation, expressive language + articulation, and receptive language. The system relies on an implementation of the Partition Around Medoids (PAM) algorithm to generate clusters of subject profiles with two levels of granularity, first considering broad diagnosis terms and medical conditions, and then looking at the specific communication skills affected. The proposal has been tested in collaboration with expert pathologists from three special education institutions of Ecuador, who were about 90% satisfied with the quality of the therapy plans provided. It was found that the two-level clustering is a crucial feature to tell apart individuals who have similar speech–language limitations, but arising from different medical conditions and, therefore, requiring different treatment.



## 1. Introduction

The acquisition of language and communication abilities is one of the most important mainstays of the brain development of humans. It is crucial for every individual to attain the tools to express needs, to learn, to be related with the environment, and in general, to have the opportunity to develop him/herself as an active member of society. According to the National Institute of Deafness and Other Communication Disorders (2014), nowadays 15 million people suffer from stutter in the world, whereas around 6 million in the United States have language impairments, and one of every 10 Americans has experienced or lived with some type of communication disorder. In Ecuador, 3% of the people suffer from some type of disability and, among them, 13% suffer disorders related with speech, hearing and language (Consejo Nacional de Discapacidades, 2014). The problem is more complicated in the ambit of special education, because around 70% of children with disabilities face learning problems derived from speech and language disorders (Parrilla & Sierra, 2013).

The aim of Speech and Language Therapy (SLT) is to maximize the people's ability to communicate, through speech, gesture and supplementary means. Speech and Language Pathologists (SLPs) assess, diagnose and provide treatment for a wide spectrum of communication disorders, which may be associated with other pathologies (e.g. cerebral palsy, autism, etc) or not (Pennington, Goldbart, & Marshall, 2003). Table 1 lists some of the main disorders as per the classifications provided by Aronson and Bless (2009, chap. 4) and Damico, Müller, and Ball (2010). The main subclassifications are also included (e.g. the hearing loss can be conductive, sensorineural or mixed) but the deepest subclassifications (e.g. differentiating palatal, labial and other types of dysglossia) are omitted.

Commonly, an SLP works with his/her patients on a weekly basis. The SLT sessions go on for about an hour, during which the SLP works with the patient in various areas, conducting exercises from basic level (breathing, swallowing, tongue control, etc) to advanced level (sentence construction, execution of complex orders, etc) and registering the patient's performance and progress. At the end of each session, the SLP is expected to generate (or update) a personalized therapy plan, borrowing elements from an ever-growing (and evolving) set of hundreds or thousands of activities. The problem we address in this paper is that, going through a routine of working with several subjects a day, SLPs

* Corresponding author. Tel.: +34 986 813967.
*E-mail addresses:* vrobles@ups.edu.ec (V.E. Robles-Bykbaev), mlnores@det.uvigo.es (M. López-Nores), jose@det.uvigo.es (J.J. Pazos-Arias), darevalol@est.ups.edu.ec (D. Arévalo-Lucero).

http://dx.doi.org/10.1016/j.eswa.2015.06.011
0957-4174/

**Table 1**
Some of most common speech–language related disorders, along with the corresponding ICD-10-CM classification codes if available (American Speech-Language-Hearing Association, 2014).

| | | | | | |
|---|---|---|---|---|---|
| Language disorders | Delay in speech and language acquisition | Mild (**F70**)<br>Moderate (**F71**)<br>Severe (**F72**) | Hearing disorders | Hearing loss (**F80.4**) | Conductive<br>Sensorineural<br>Mixed |
| | | Deep (**F73**) | | Deafness (**H90**) | |
| | Specific developmental disorders of speech and language | Expressive language disorder (**F80.1**) | Speech disorders | Articulation disorders (**F80**) | Phonological |
| | | Mixed receptive–expressive language disorder (**F80.2**) | | | Dysglossia |
| | Aphasia (**R47.0**) | Expressive (Broca) | | | Dysarthria |
| | | Sensory (Wernicke) | | Fluency disorders | Slurred speech (**F47.81**) |
| | | Conduction | | | Stuttering (**F80.81**) |
| | | Transcortical sensory | | | Childhood onset fluency disorder (**F98.5**) |
| | | Transcortical motor | | | Disfluency Fluency and rhythm disorders |
| | | Mixed Transcortical | | Voice disorders | Dysphonia (**R49.0**) |
| | | Anomic | | | Aphonia (**R49.1**) |
| | | Global | | | Hypernasality and hyponasality (**R49.2**) |
| Swallowing disorders | Aphagia (**R13.0**)<br>Dysphagia (**R13.1**) | | Communication disorders | Organic mutism Selective mutism (**F94.0**) | |

hardly ever manage to keep in mind the whole set of activities they can put into the therapy plans. As a result, they end up using only a subset of activities, implying that the therapy plans they make are often sub-optimal.

The aim of our research is to develop an expert system to aid the SLPs in the generation and updating of therapy plans, considering activities from five areas: hearing, oral structure and function (oral-peripheral mechanism), linguistic formulation, expressive language + articulation, and receptive language. Specifically, the system – which we have called SPELTA (*SPEech and Language Therapy Assistant*) – was devised to automatically generate two kinds of therapy plans:

- A **master plan** indicates the general activities (from the five aforementioned areas) that must be conducted with each patient during a period of six months. Table 2 contains an example of a child's profile and a master plan designed by an SLP, indicating a number of activities to improve some affected skills. One example of such general activities could be "*perform blow exercises to increase the blow force*". The set of activities defined for a certain speech–language area make up a **subplan**.
- A **specific plan** indicates daily exercises for the six-month period, distributed over a certain number of therapy sessions, each one lasting for 1 to 2 h. The exercises are chosen according to the general activities of a master plan on the grounds of diversity, duration, strain and other parameters. Two specific exercises related to the abovementioned activity could be "*blow confetti 10 times during 2 s*" or "*inflate one ballon in no more than 6 exhalations*".

In this paper we present the general architecture of the SPELTA system and the procedure designed to perform the automatic generation of master plans for speech–language therapy, which uses a custom version of the PAM (*Partition Around Medoids*) algorithm to generate clusters of subject profiles with two levels of granularity, first considering broad diagnosis terms and then looking at the specific affected skills. This system has been put to the test in a pilot experiment conducted with the aid and supervision of a team of expert SLPs who are treating children in three institutions of special education in Ecuador: Instituto de Parálisis Cerebral del Azuay (Institute of Cerebral Palsy of Azuay), Fundación "General Dávalos" (General Davalos Foundation) and CEDEI School.

Next, we give an overview of previous works in computer-aided SLT (Section 2). After that, Section 3 presents the SLT environment in which the SPELTA system has been implemented, along with the details behind the automatic generation of master therapy plans. The experiments we have done hitherto are described in Section 4. Conclusions and future work are finally given in Section 5.

## 2. Related work of ICT support and expert systems for SLT

Over the last decade, there have been several approaches to apply information and communication technologies (ICT) to support different processes of SLT, including many applications of expert systems. The proposals typically focus on one of the major stages: diagnosis, therapy design (which is the aim of our work) or therapy enforcement.

### 2.1. Diagnosis

During the last few years, several authors have worked to automate diagnosis tests by means of audiovisual signal processing. For example, Schipor, Pentiuc, and Schipor (2012) presented a model for automatic assessment of pronunciation quality for children, using *hidden Markov models* (HMM) and implementing a correlation measure to compare the level of intelligibility of new utterances presented to the system, whereas Saz et al. (2009) had used HMM in combination with a subword-based pronunciation verification method. Utianski, Sandoval, Lehrer, Berisha, and Liss (2013) developed an application able to record speech samples and provide a set of derived calculations with the aim of assessing the integrity of speech production (vowel space area, assessment of an individual's pathology fingerprint, and identification of parameters of the intelligibility disorder). In the same line of research, Caballero-Morales and Trujillo-Romero (2014) improved the recognition rates for dysarthric patients by integrating multiple pronunciation patterns in an expert system using genetic algorithms. Later on, Mustafa, Rosdi, Salim, and Mughal (2015) provided a thorough analysis of general and specific factors that affect the recognition accuracy of that system and previous ones.

**Table 2**
A sample master plan designed by a Speech and Language Pathologist for a real case.

| Case | 47 |
|---|---|
| Profile resume description | *Age:* 8 years, 5 months and 5 days<br>*Medical diagnostic:* High-functioning autism. Hyperactivity. Attention deficit. (ICD-10-CM codes **F72, F90.1**)<br>*Speech and language diagnostic:* Dysarthria (ICD-10-CM code **I69.222**) |
| *Therapy plan designed by SLP* | |
| Area | Exercises/Areas to work |
| Hearing | Discrimination of sounds of nature, body and animals<br>Detect presence or absence of sounds<br>Working on rhythmic structures using loud and soft sounds<br>Working with simple rhythmic sequences<br>Discrimination of similar phonetic sounds |
| Oral structure and function | Performing massages to stimulate the phono-articulatory apparatus<br>Tongue exercises: perform slow and fast movements, move to<br>left and right, and stick the tongue out<br>Performing exercises with lips (retraction and protrusion)<br>Performing passive exercises of the phono-articulatory apparatus |
| Linguistic formulation | Performing relaxation exercises for intra and extra-laryngeal musculature<br>Working with inspiration–expiration long and short exercises<br>Performing blowing-exercises without supporting material<br>Performing blowing-exercises using supporting material |
| Expressive language + articulation | Constructing sentences from a given word<br>Sorting the words of a sentence<br>Working in grammatical structure<br>Developing the spontaneous conversation<br>Performing activities that use twisters and rhymes<br>Working with the personal articulation exercise book |
| Receptive language | Working with sequences of 3–4 pictograms, ordered from left to right<br>Identifying objects on the right side of the body<br>Dealing with directional commands<br>Solving remote situations<br>Logical reasoning with numbers from 1 to 10 |

### 2.2. Therapy design

The automatic generation of therapy plans has very few precedents in the literature, even though there have been various approaches to the enabling task of classifying subject cases based on diagnosis data. For example, Hariharan, Chee, Chia Ai, and Yaacob (2012) dealt with the classification of speech dysfluencies using spectral features, trying to aid the SLPs in what has been traditionally a subjective, inconsistent, time consuming and error-prone task. Likewise, Verikas, Bacauskiene, Gelzinis, Vaiciukynas, and Uloza (2012) conducted a comparative study of different voice analysis systems, in order to classify healthy patients and those suffering from diffuse laryngeal defects. The classification is done by grouping a set of cases (data elements) into groups (clusters), so that cases within the same group are similar but cases in different groups are dissimilar. Some of the algorithms that can be used for this task are surveyed in Frades and Rune (2010).

Among the works that did attempt to automate the design of therapy plans, we can highlight the work of Schipor, Pentiuc, and Schipor (2010), who developed an expert system based on fuzzy logic to plan sessions for dyslalia treatment. Their system uses three types of information to define the inference rules: (i) social, cognitive and affective parameters, (ii) homework reports, and (iii) test scores. With the inference rules, the system provides outputs about the frequency, duration and type of exercises of the therapy sessions. Later on, Yeh, Hou, and Chang (2012) presented an approach that used artificial neural networks (ANN) for the classification of subject profiles along 127 attributes, and thereupon applied classification and regression tree (CART) techniques to assist pathologists for precise assessment and appropriate treatment of a wide range of occupational therapy problems (which may be seen as a superset of speech–language problems). This wide scope makes the work of Yeh, Hou & Chang most similar to ours; for that reason, in Section 4 we will compare the performance of our PAM-based approach and the performance achieved by their use of ANN and CART.

### 2.3. Therapy enforcement

Many research efforts have separately shown that ICTs have great potential to improve the enforcement of speech–language therapy plans, extending the continuum of care and enabling better clinical outcomes. Mashima and Doarn (2008) compiled the outcomes of 40 studies with telehealth models used by SLPs to provide services to patients with various cognitive-communication disorders: aphasia, dysarthria, apraxia, dementia resulting from cerebrovascular disease, traumatic brain injury, Parkinson's disease, cerebral palsy and multiple sclerosis. Nonetheless, the technology employed in those studies (training videos, recorded speech samples, e-mail, audio- and video-conference) fell short of automating any of the SLT processes, demanding heavy involvement of SLPs and caregivers. More recently, Palmer et al. (2012) argued that the advances in artificial intelligence and audiovisual signal processing make it possible to automate certain SLT tasks and thus deliver more intensive and efficient therapy out of the clinics, while making better use of the therapists' time. The state-of-the-art in this area includes web-based developments (Ooi, Raja, Sung, Fung, & Koh, 2012), mobile applications (Bunnell & Beidel, 2013) and various types of robots (Choe, Jung, Baird, & Grupen, 2013; Kose, Akalin, & Uluer, 2014), featuring components of voice-based and gesture-based interaction (Hogrefe, Ziegler, Wiesmayer, Weidinger, & Goldenberg, 2013; Sekine & Rose, 2013), avatars in the role of 'virtual therapists' (Abad et al., 2013; Teodoro, Martin, Keshner, Shi, & Rudnicky, 2013), etc. Expert systems have been used in this area, for example, to adapt parameters

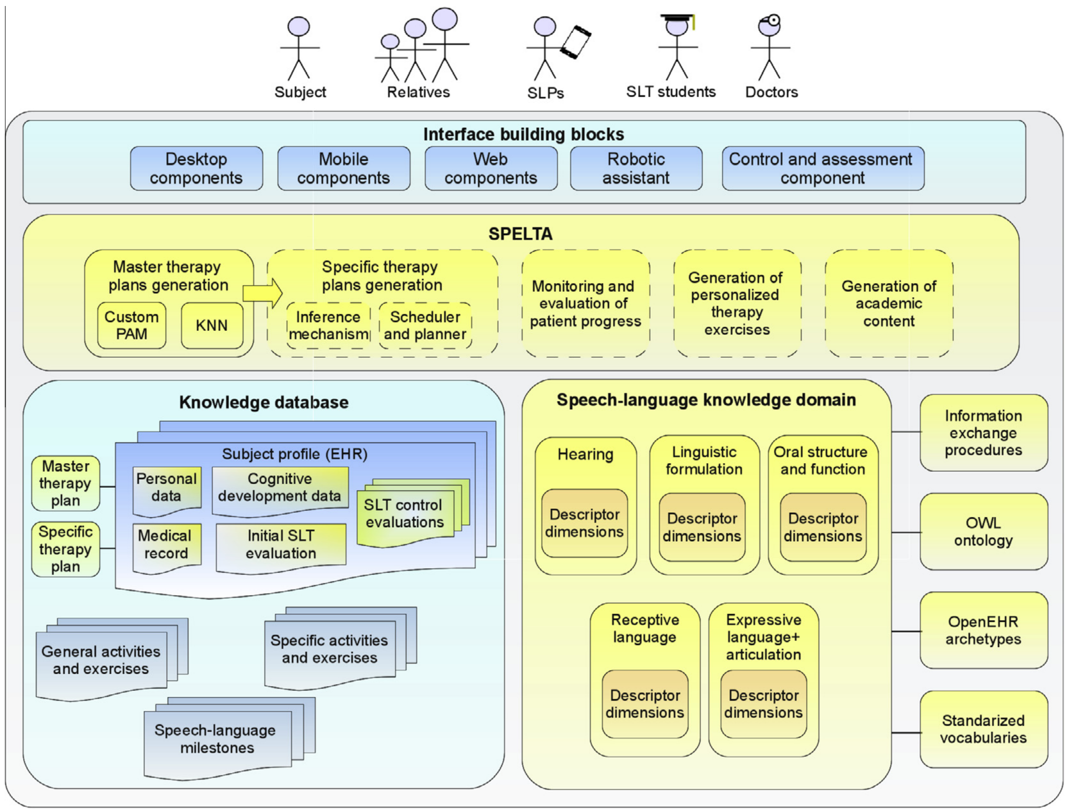


**Fig. 1.** The high-level design and main components of our SLT environment.

of the user interfaces, in order to make them more amenable and engaging (Kostoulas et al., 2012).

## 3. The SPELTA system: Overview and procedures involved in therapy plan generation

The SPELTA system is one part of an ecosystem of ICT tools designed from the ground up to support SLT within a fully integrative environment for clinicians and students, pathologists, patients, relatives and other potential users. The high-level design of the environment is depicted in Fig. 1, with some greater detail for the modules that are relevant for this paper.

Our environment is based on a formal knowledge model of the SLT domain, which includes an OWL ontology of concepts and instances expressed through OpenEHR archetypes – see Robles-Bykbaev, López-Nores, Pazos-Arias, García-Duque, and Ochoa-Zambrano (2014) for further details. This model provides the foundations for the following tasks, among others:

- Accessing, sharing and querying the information according to specialized taxonomies of SLT concepts and user types.
- Automating statistical procedures to analyze the patients' evolution, the effectiveness of the applied therapies, common SLT patterns, behavioral patterns, etc.
- Automating the adaptation of contents to put in therapy plans or learning courses, according to SLT taxonomies and patient/student profiles.
- Integrating assistive technologies to provide support during the therapy sessions: robotic assistants, mobile applications, remote monitoring, etc.
- Developing inference mechanisms for recommender and decision-support systems to assist in the preparation of therapy plans, the evaluation of exercise results, the generation of case studies, etc.
- Porting the data structures through different architectures and systems.

The different systems work with a knowledge base that comprises four main data structures: the subject profiles (in the shape of OpenEHR health records), a library of general activities to put in the master therapy plans, a library of exercises for the specific plans and a list of speech–language milestones and skills according to chronological and developmental age (e.g. formal expressions of facts like "*from 1 to 2 years of normal development, a child must to be able to acquire new words on a regular basis, know a few parts of the body and point to them when asked*"). For the purposes of this paper, it is particularly important to look into the subject profiles, which contain the following items:

- **Personal data**, including chronological age, gender, name, etc.
- A **medical record**, specifying diagnosis, general medical conditions and related disorders (e.g. cerebral palsy, hemiparesis, athetosis, etc) as indicated by doctors.
- A **record of cognitive development data**, indicating cognitive age, gap in language development, expressive language age and receptive language age, as estimated by SLPs. These data are expressed in units of years.
- An **initial SLT evaluation** that, in the current state of implementation, looks at 102 dimensions from the subareas indicated in Table 3.[1]
- **Control evaluations** with the results of successive therapy sessions.

[1] For example, dimension 1 is the part of hearing evaluation that assesses the subject's capability to locate sound sources and voice response, whereas dimension 63 is a test from the Preschool Language Scale test (PLS) that evaluates whether the subject can imitate two different words or sounds.

**Table 3**
The main tests and areas for the initial evaluation of subjects conducted by SLPs with the current version of the SPELTA system.

| Test | Areas of evaluation | Test | Areas of evaluation |
|---|---|---|---|
| Hearing | Reflex<br>Localization of sound sources<br>Response to voice | PLS (communication subscale)<br>and articulation test | ***Expressive language***: vocal development, social communication, semantics (content)-vocabulary and concepts, structure (form)-morphology and syntax, and integrative thinking skills |
| Oral function and structure | Tongue<br>Teeth<br>Palate<br>Lips | | ***Articulation***: phonemes, sentences, polysyllabic words, vowel diphones |
| Linguistic formulation | Maxillary mobility<br>Phonation<br>Breathing condition | PLS (auditory comprehension subscale) | ***Receptive language***: attention, semantics (context)-vocabulary and concepts, structure (form)-morphology and syntax, and integration skills. |

- Links to **master and specific therapy plans**, either designed by SLPs or generated automatically by the SPELTA system.

For the generation of therapy plans, the SPELTA system is mainly concerned with the topmost four of the aforementioned items. The subject profiles available become elements of a corpus by having the data of the medical record and the initial SLT evaluation encoded into binary form. Fig. 2 shows an example of the binary encoding of three dimensions corresponding to the evaluation of voice quality. One of the dimensions is represented by one bit only, with '0' meaning "*wrong response*" to a given test and '1' meaning "*right response*". Other aspects are represented by more bits, in such a way that (whenever it makes sense) the number of different bits defines an order relation. For example, representing the intensity of voice production by '00' when it is "*weak*", by '01' when it is "*moderate*" and by '11' when it is "*strong*", we convey the meaning that "*moderate*" is halfway between the others, which is important for the similarity metrics used in the clustering processes (further details below).

Following the binary encoding, as shown in Fig. 3, the subject cases are organized into clusters with two different levels of granularity, separately for each one of the five speech–language areas. We chose the PAM algorithm for the SPELTA system because it can work with cases that are not only represented by numbers (the SLT subject profiles contain non-numeric data) and because each one of the five speech–language areas involves different computations in the similarity metrics. Likewise, with the PAM algorithm we do not need to define the density-threshold that is required in other clustering algorithms, given that it cannot be calculated a priori. As explained in Frades and Rune (2010),

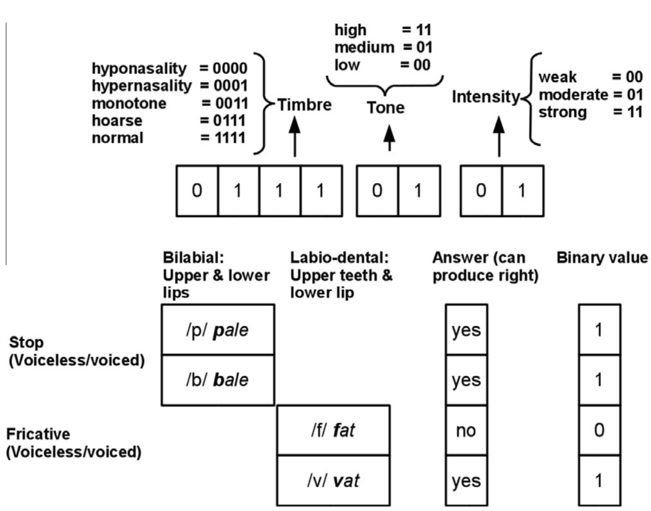

**Fig. 2.** Sample binary coding of some features of voice quality and articulation.

PAM proceeds by searching for the K representative cases (the medoids) that yield the minimum average dissimilarity within each cluster according to whichever metrics; after finding the K medoids, K clusters are constructed by assigning each case to the nearest medoid. Our advisory team of SLPs deemed it important for our system that each representative element of the clusters would be a real case, in order to facilitate inspection of the corpus of SLT subjects. Having said this, the clustering process works as follows:

In the first level, we aim to separate subjects who may have similar speech–language skills and limitations, but arising from such different illnesses as mild cerebral palsy (which allows an individual to move without assistance) and spastic quadriplegia (by which all four limbs are paralyzed). Different conditions and different sources of the cognition deficits have to be treated with different general activities in the master plans and different exercises in the specific plans. Accordingly, the subjects are clustered as per the information stored in the medical record of their profiles. The PAM algorithm is driven by the distance metrics of Eq. (1), where $S_i$ and $S_j$ refer to two different subjects, $A$ is one of the speech–language areas, $f$ goes over the set of features from the medical record which are relevant for that area ($features_{MR}(A)$), and ManhDist denotes the mean-Manhattan binary distance (Khalifa, Haranczyk, & Holliday, 2009).

$$d_1(S_i, S_j, A) = \sum_{f \in features_{MR}(A)} \text{ManhDist}(f(S_i), f(S_j)) \quad (1)$$

In the second level, we cluster the subjects as per the fine-grained evaluation of the speech–language skills and impairments, the record of cognitive development data and the results of the initial SLT evaluation. For example, within a first-level cluster that includes the cases of children with Down syndrome and phonological disorders, we need to differentiate subjects who commit additions (adding extra sounds in some words, e.g. "*balue*" for "*blue*") from subjects who commit substitutions (one or more sounds are substituted for another, e.g. "*bagon*" for "*wagon*"). To this aim, we run the PAM algorithm to organize the subjects included in each one of the first-level clusters, this time driven by the distance metrics of Eq. (2). The first summation measures the mean-Manhattan binary distance of the initial SLT evaluations of two subjects, considering only the dimensions which are relevant to the speech–language area in question, $dimensions_{IE}(A)$. The second summation provides a scale factor derived from the absolute differences of cognitive age, gap in language development, expressive language age and receptive language age (the features of cognitive development data, $CDD$).

$$d_2(S_i, S_j, A) = \sum_{d \in dimensions_{IE}(A)} \text{ManhDist}(d(S_i), d(S_j)) \cdot \sum_{f \in features_{CDD}(A)} |f(S_i) - f(S_j)| \quad (2)$$

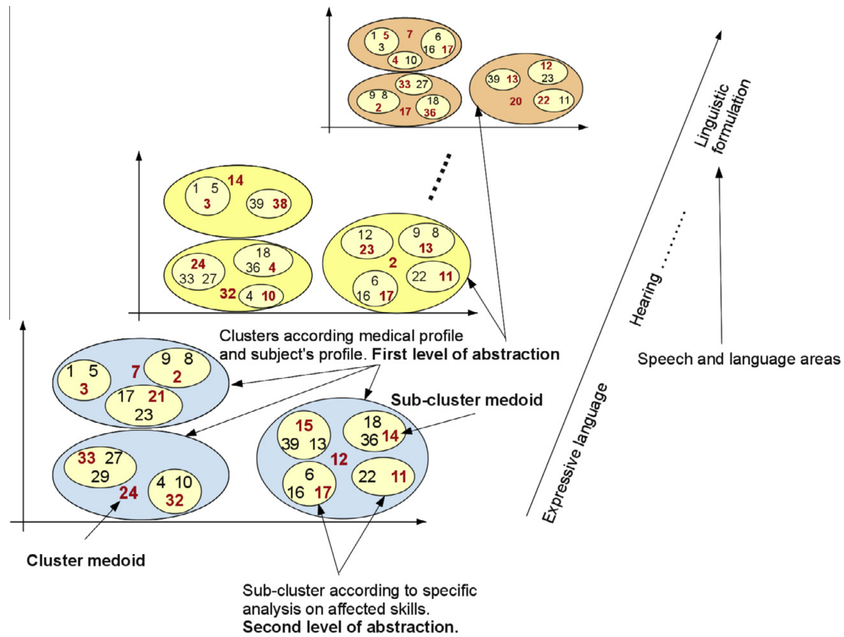


**Fig. 3.** The clustering approach of the SPELTA system.

Each one of the first-level and second-level clusters has one of the subject cases designated as a medoid. This facilitates the classification of new cases, identifying the closest subjects in each one of the speech–language areas. We can identify the first-level cluster for a new case simply by looking for the first-level medoid that is closest to it according to the distance metrics of Eq. (1). Then, analogously, we select the second-level cluster for the new subject by using Eq. (2) to measure its distances with respect to the second-level medoids.

Having performed the classification of a new subject, the SPELTA system finally composes the master therapy plan for it from the subplans defined for the subjects who were found to be closest (i.e. the most similar cases) in each one of the speech–language areas. The new plan is presented to the SLPs through visual interfaces, so that they can validate it as a whole or modify certain parts, as they deem necessary. In order to facilitate the SLPs' task, the interfaces display which cases the activities were borrowed from and, if several known subjects were found to be equally distant to the new one in some of the areas, it is possible to browse the superset of activities, the intersections and the disjunctions. For example, Table 4 shows a real master plan generated by the SPELTA system, with the third column indicating the most similar subjects in each area and the features that make them similar to the new case.

Through the aforementioned procedure, a master therapy plan for a new subject can be derived from a number of master plans already included in the corpus of the SPELTA system. The value and diversity of these automatically-generated plans will be greater if the corpus is progressively augmented with plans contributed by a number of SLPs, as it happens with the deployment of our solutions among the special education institutions of Ecuador. The procedures for information exchange enabled by the formal knowledge model presented (Robles-Bykbaev et al., 2014) are therefore a crucial feature for our system.

## 4. Experiments and results

In order to assess the value of the automatic generation of master therapy plans for the SLP's work, we have conducted a pilot experiment with the collaboration of three institutions of special education in Ecuador: Instituto de Parálisis Cerebral del Azuay (Institute of Cerebral Palsy of Azuay), Fundación "General Dávalos" (General Davalos Foundation) and CEDEI School. A team of 6 experts was provided with an online tool to conduct evaluations of 53 children with different types of disabilities and cognitive ages from 0 to 7 years. The tool allowed them to browse the taxonomy of speech–language disorders and the corpus of subject cases (Fig. 4), to conduct the initial evaluation of the 102 dimensions mentioned in Section 3 for a new case (Fig. 5) and to examine the activities and exercises of the therapy plans (Fig. 6).

The SLPs created 40 master plans, each one indicating activities suited to the particular skills and disabilities of one child in the five speech–language areas. Those 40 subject profiles were put into a corpus and the clustering algorithms were run. The cases of the remaining 13 children were presented to the SPELTA system to automatically compose master plans out of the set of 200 ($40 \times 5$) subplans. The SLPs would then debate whether each one of the plans was convenient or not, considering its accurateness, consistency and completeness. Then, by consensus, they would rate each subplan with a rating of 1 (good) or 0 (bad). The results are summed up in Table 5, showing that the system attained very good marks (an average rating of 0.908), with perfect ratings in three of the areas.

In order to get more insight into the convenience of the clustering procedure and the factors of the metrics presented in Section 3, we asked the SLPs to assess the master plans obtained through slightly different approaches:

(1) The first alternative implied no clustering, but rather it applied only the first summation of Eq. (2) to identify, from among all the cases in the corpus, the most similar cases in each speech–language area judging from the outcomes of the tests of the initial SLT evaluations. The master plans were composed out of the subplans attached to those cases in the corresponding area. Data from the medical record and the record of cognitive development data were not considered.
(2) The second alternative did not compute any clusters either, but looked for the most similar cases in each area considering the information stored in the medical record, in the

**Table 4**
A master plan generated by the SPELTA system, with indications of the subjects from which the activities of the different speech–language areas were borrowed.

| Case | 52 | |
|---|---|---|
| Profile resume description | *Age:* 15 years and 8 months<br>*Medical diagnostic:* Athetoid cerebral palsy (ICD-10-CM code **G80.3**)<br>*Speech and language diagnostic:* Mixed receptive–expressive language disorder (ICD-10-CM code **F80.1**) | *Receptive language age:* 4 years<br>*Expressive language age:* 2 years, 8 months<br>*Language developmental age:* 3 years and 4 months |
| *Therapy plan generated by the SPELTA system* | | |
| Area | Exercises/Areas to work | Source subplan(s) |
| Hearing | Perform exercises to sounds identification<br>Discriminate of sounds of nature, body and animals<br>Perform phonemes discrimination exercises | **Case 37:** a patient with a similar receptive language age (4 years and 6 months) and a 100% coincidence in the evaluation of hearing (cochleopalpebral reflex, startle response, turns head to sound source, identifying sound objects, sound source localization without visual stimulus) |
| Oral structure and function | Perform segmental relaxation massages<br>Perform slow and fast tongue movements<br>Perform exercises with lips (retraction and protrusion)<br>Achieve sound productions using the oropharynx structure<br>Perform active and passive exercises using tongue, lips and jaw | **Case 18:** a patient with an 84% coincidence in the oral peripheral mechanism (same tongue size, same speed in tongue movements, present tongue protrusion, voluntary and involuntary swallowing are present, is able to chewing hard and soft food, sialorrhea is not present) |
| Linguistic formulation | Work in the automatic respiration process (inspirations and expirations)<br>Work with blow exercises to increase the blow force | **Case 22:** a patient with a 70% coincidence in linguistic formulation (same respiratory frequency, same thorax symmetry, diaphragmatic breathing) |
| | Respiration exercises associated to vowels and simple phonemes (/pa/, /da/, /fo/) | **Case 3:** a patient with a 70% coincidence in linguistic formulation (diaphragmatic breathing, no nasal obstruction, same exhalation period) |
| Expressive language + articulation | Construct sentences from a given word<br>Sort out the words of a sentence<br>Work in grammatical structure<br>Develop the spontaneous conversation<br>Perform activities that use twisters and rhymes<br>Work with the personal articulation exercise book | **Case 22:** a patient with a similar expressive language age (1 year and 7 months), a similar diagnose for the medical examination (cerebral palsy and mixed receptive–expressive language disorder) and a 100% coincidence in the speech–language evaluation |
| Receptive language | Work with sequences and puzzles of 4 elements<br>Learn semantic categories<br>Identify objects according to their utility<br>Identify daily activities<br>Learn temporal notions (day and night, before and after)<br>Identify similar/distinct objects according to their utility | **Case 37:** a patient with a similar receptive language age (4 years and 6 months), similar diagnoses for the medical examination (cerebral palsy and mixed receptive–expressive language disorder) and a 90% coincidence in the speech–language evaluation (the only difference relates to the use of place prepositions like "under", "over", etc) |

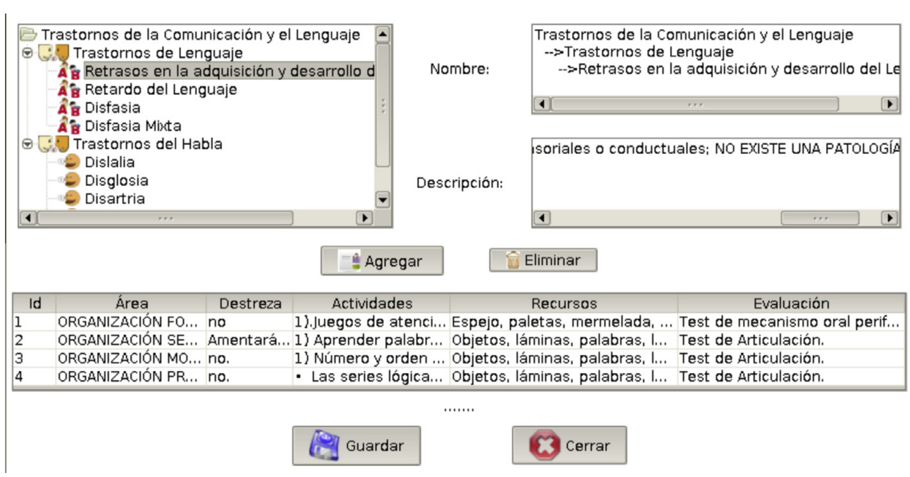


**Fig. 4.** Snapshots of our online tool for SLPs: Browsing the taxonomy and descriptions of speech–language disorders.

record of cognitive development data and in the initial SLT evaluation, all in one go through the distance metrics of Eq. (3):

$$d_{Alt2}(S_i, S_j, A) = \sum_{f \in features_{MR}(A)} \text{ManhDist}(f(S_i), f(S_j)) \cdot \sum_{d \in dimensions_{IE}(A)} \text{ManhDist}(d(S_i), d(S_j)) \cdot \sum_{f \in features_{CDD}(A)} |f(S_i) - f(S_j)| \quad (3)$$

(3) In the third alternative, we used the PAM algorithm to generate one-level clusters only, driven by the initial SLT evaluations (i.e. applying the first summation of Eq. (2) only). Again, data from the medical record and the record of cognitive development data were not considered. The master plans were composed out of the subplans attached to the most similar medoids in each speech–language area.

(4) In the fourth alternative, we also use the PAM algorithm to generate one-level clusters, but this time driven by the

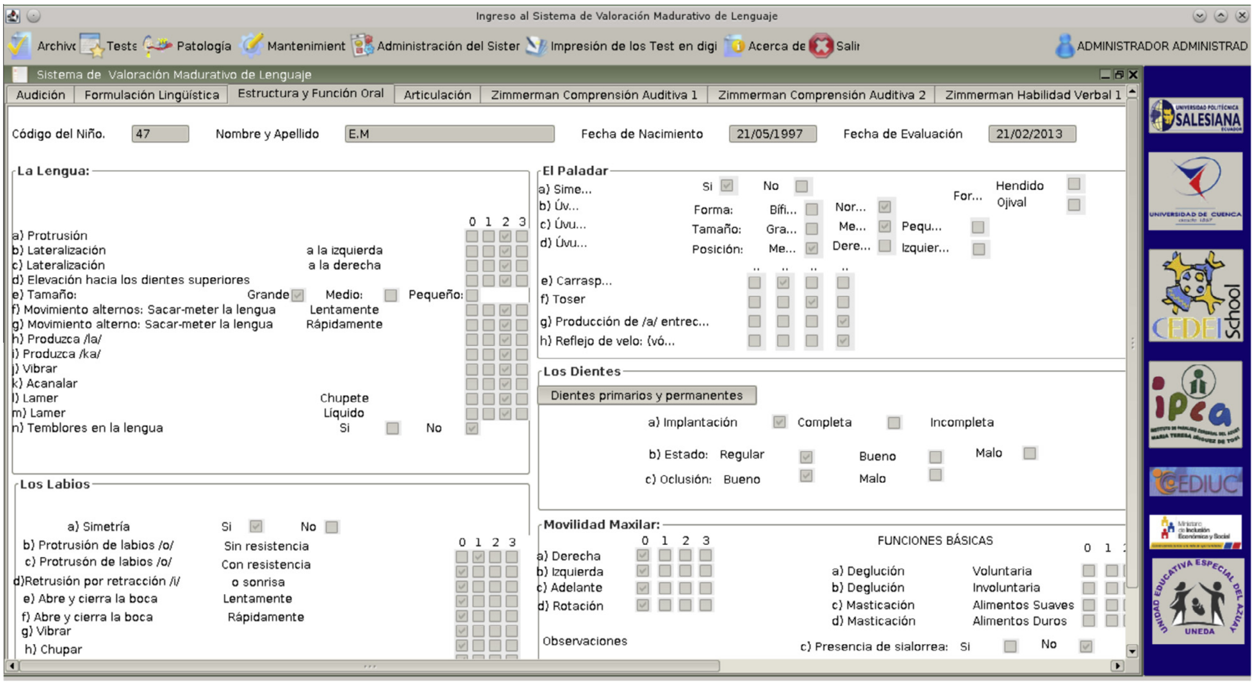


**Fig. 5.** Snapshots of our online tool for SLPs: Evaluating the dimensions of oral structure and function.

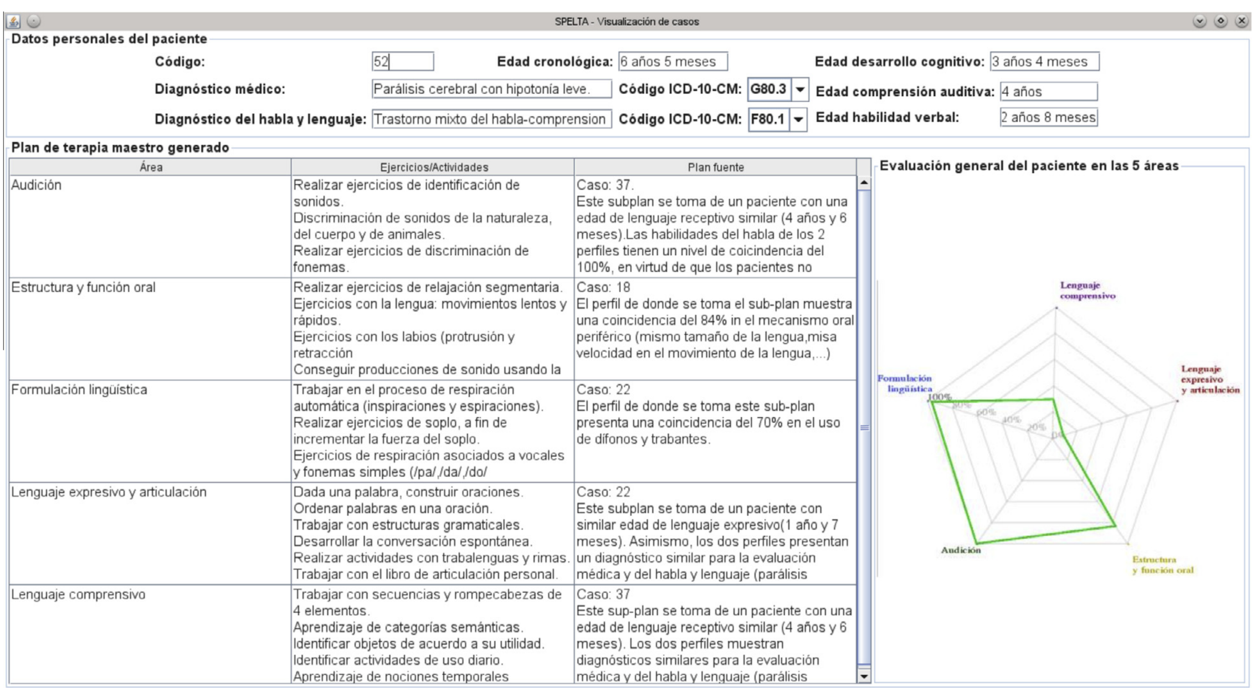


**Fig. 6.** Snapshots of our online tool for SLPs: Examining a master therapy plan.

**Table 5**
Rating of the SPELTA system in the generation of master therapy plans.

| Speech–language area | Rating |
|---|---|
| Hearing | 1 |
| Linguistic formulation | 1 |
| Oral structure and function | 1 |
| Expressive language + articulation | 0.692 |
| Receptive language | 0.846 |
| Average | 0.908 |

**Table 6**
Average ratings achieved by the SPELTA system and the alternatives approaches.

| Approach | Rating |
|---|---|
| SPELTA | 0.908 |
| Alternative 1 | 0.231 |
| Alternative 2 | 0.477 |
| Alternative 3 | 0.677 |
| Alternative 4 | 0.200 |
| ANN + CART | 0.820 |

**Table 7**
Ratings obtained by SPELTA and the alternative approaches in the evaluation of 13 master plans using a Likert scale.

| Approach | Hearing | Oral structure and function | Linguistic formulation | Expressive language + articulation | Receptive language |
|---|---|---|---|---|---|
| SPELTA | **63** | **63** | **64** | 43 | **59** |
| Alternative 1 | 41 | 43 | 44 | 38 | 37 |
| Alternative 2 | 43 | 47 | 50 | 38 | 44 |
| Alternative 3 | 59 | 62 | **64** | 37 | 36 |
| Alternative 4 | 40 | 42 | 44 | 27 | 36 |
| ANN + CART | 55 | 56 | 52 | **54** | 55 |

The greatest value in each SLT area is highlighted in bold.

distance metrics of Eq. (3), i.e. taking into account the information stored in the medical record, in the record of cognitive development data and in the initial SLT evaluation.

(5) Finally, we replicated the approach of Yeh et al. (2012) involving artificial neural networks (ANN) and classification and regression trees (CART), with minor adaptations to deal with the data of our subject profiles (which capture only aspects of speech and language skills, but with greater detail and granularity than the models of Yeh et al. for occupational therapy). The algorithms here also proceed in two stages to deal with the critical attributes first, and with the important ones next.

Having evaluated the master plans provided by SPELTA and the five alternative approaches in the same conditions, we got the average ratings shown in Table 6, which show that the two-level clustering of SPELTA is the right way to go in order to put together subjects not only because their speech–language skills and limitations are similar, but also (and more importantly) because those limitations have similar origins or similar concurrent disabilities, and therefore can be treated with similar activities and exercises. Moreover, the two-level clustering helped the SLPs to discover "*hidden populations*", with groups of subjects that suffer from the same disorders and have the same cognitive development age, but have different impairments in the same speech–language area.[2]

The results attained by the alternatives, in contrast, point out some inconsistencies in their design due to the unclear effects of applying one-level clustering or not, and dealing with more or less information from the subject profiles. For example, alternatives 1 and 3 worked with the same data and applied the same distance metrics, but the outputs of alternative 3 (with one-level clustering) were much better than those of alternative 1 (no clustering), whereas alternative 2 (no clustering) attained significantly better ratings than alternative 4 (one-level clustering). The approach of Yeh et al. (2012) involving neural networks and regression trees attained better results than these alternatives, though it was found to be around 8% less precise than the two runs of PAM we implemented in the SPELTA system. analyzing the computations of the ANN + CART approach with the aid of the collaborating SLPs, we found that the hierarchical structure of the decision trees was too rigid, not allowing to reconsider decisions driven by medical conditions in light of any particular observations related to the communication skills. The SPELTA approach is more flexible in this regard.

Looking for a more fine-grained evaluation, we asked the team of SLPs to rate (again, by consensus) each one of the subplans provided by the different approaches for each one of the speech–language areas using the Likert scale, i.e. using 1 (totally disagree), 2 (disagree), 3 (neutral), 4 (agree) and 5 (totally agree). This way, since there were 13 master plans to evaluate, each approach could get from 13 to 65 points in each area. The results, shown in Table 7, reveal that SPELTA provided the most convenient activities in all the areas, except for the area of expressive language + articulation, in which the ANN + CART approach of Yeh et al. (2012) turned out to be the best. It is worth noting ANN + CART attained the most uniform ratings across the different areas, which may suggest that its algorithms do not depend much on the granularity of the modeling of the respective phenomena. In contrast, the SPELTA system achieved moderate results in expressive language + articulation, whereas it behaved very well in the others. This disparity suggests that the performance of our algorithms in that specific area may be improved by refining the number of parameters considered and the level of detail in their modeling.

## 5. Conclusions and future work

We have presented an expert system that generates therapy plans for people with speech–language disorders, handling medical data, cognitive development data and the results of 102 tests of speech–language skills. The SPELTA system is intended as an aid for pathologists to identify the most suitable activities for each subject, reusing subplans from other subjects stored in a corpus. SPELTA has attained very satisfactory results in a pilot experiment conducted with the aid and supervision of a team of expert speech–language pathologists from three institutions of special education in Ecuador, who provided positive ratings for 13 master therapy plans derived automatically from other 40 plans used for training. The evaluation results show that it is a convenient approach to classify the subject cases in two rounds, in order to tell apart individuals who have similar speech–language limitations, but arising from different medical conditions and, therefore, requiring different treatment.

The most relevant precedent to our approach was the expert system of Yeh et al. (2012), which used artificial neural networks and classification and regression trees for a different purpose, namely the design of occupational therapy plans for children. That approach proceeds in two rounds, too, but the hierarchical structure of the decision trees turned out to be somewhat inflexible for the SLT domain, since it prevents from reconsidering decisions suggested by medical conditions in the light of any particular observations related to the communication skills. In our experiments, however, neural networks and decision trees attained more uniform performance across different areas of SLT than our approach, which seems to be more sensitive to the number and granularity of diagnosis parameters considered. It is well known that neural networks are especially advantageous for working well with data that contains noise, has a poorly understood structure or changing characteristics (Hussin, Kamel, & Nagi, 2004).

Other conceptual precedents to the design of our expert system can be found in the area of document retrieval. For example, Zhang and Chow (2012) presented an approach that represented documents by a two-level structure (document level and paragraph

[2] One such hidden population involved a group of children suffering from spastic hemiplegic cerebral palsy (code G80.2), some of which had certain articulation skills much most deteriorated than other subjects with the same cognitive age and the same illnesses.

level) to capture global and local semantics, and then dealt with the matching between documents and queries as an optimization problem. Two-level clustering can also be found in works aiming to detect topics and produce highlights in Twitter (Petkos, Papadopoulos, & Kompatsiaris, 2014). Multi-level clustering for large databases was considered in Lechevallier and Ciampi (2007), with an example of dealing with nutritional data from a study on nutrition and cancer. Nguyen, Phung, Nguyen, Venkatesh, and Hai Bui (2014) recently considered other case studies in various application domains – ranging from document modeling to public health – and presented a general proposal of Bayesian nonparametric multilevel clustering with group-level contexts. The contributions of this paper are focused on the area of SLT, showing that the PAM algorithm is a suitable choice for the kind of diagnosis data considered, that one only round of clustering is not enough, and that the distance metrics allow a degree of flexibility that is missing with other approaches when it comes to generate therapy plans.

Our tools have been used in the three collaborating institutions of Ecuador since April 2014, and the corpus of subject cases is growing steadily with the digitization of subject cases and therapy plans previously kept in paper form only, as well as with the introduction of new cases and plans by the pathologists. The exchange of information among the institutions is accomplished by means of formalized data structures and procedures, contributing to build a common knowledge basis that can ensure the value and diversity of the automatically-generated plans. At this time, we are working to refine the algorithms of the SPELTA system to improve the quality of the subplans generated for the areas of receptive language and expressive language + articulation (the weakest according to the results of Tables 5 and 7). To this aim, we plan to increase the granularity of the diagnosis data considered; furthermore, we want to define a new metrics to drive a third round of clustering in those areas, in order to differentiate combinations and interrelations of features that the current version of the system is treating as equivalent or proximally-related when, all in all, they demand different types of activities and exercises. Besides, in preparation for a significant growth of the corpus of subject cases, we are working to automate the decision of what the target number of clusters should be for each run of the PAM algorithm. To date, we have been using rough estimates of density in the space provided by the binary encoding of diagnosis and speech–language tests, but we have not checked whether this approach remains valid as the corpus grows.

## Acknowledgments

The authors from Universidad Politécnica Salesiana have been supported by the "Sistemas Inteligentes de Soporte a la Educación" research project (CIDII-010213). We would like to thank Zaituna Bykbaeva, Gladys Ochoa and all the collaborating people from Instituto de Parálisis Cerebral del Azuay (IPCA), Fundación "General Dávalos" and CEDEI School. The authors from the University of Vigo were 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).

## References

Abad, A., Pompili, A., Costa, A., Trancoso, I., Fonseca, J., Leal, G., et al. (2013). Automatic word naming recognition for an on-line aphasia treatment system. *Computer Speech & Language, 27*(6), 1235–1248.

Aronson, A., & Bless, D. (2009). *Clinical voice disorders* (4th ed.). New York: Thieme Medical Publishers.

American Speech-Language-Hearing Association (ASHA) (2014). *International classification of diseases, tenth revision – clinical modification, related to speech, language and swallowing disorders*. (ICD_10_CM). Retrieved October 4, 2014 from: <http://www.asha.org>.

Bunnell, B. E., & Beidel, D. C. (2013). Incorporating technology into the treatment of a 17-year-old female with selective mutism. *Clinical Case Studies, 12*(4), 291–306.

Caballero-Morales, S. O., & Trujillo-Romero, F. (2014). Evolutionary approach for integration of multiple pronunciation patterns for enhancement of dysarthric speech recognition. *Expert Systems with Applications, 41*(3), 841–852.

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.

Consejo Nacional de Discapacidades (CONADIS) (2014). *Registro nacional de discapacidades*. Retrieved May 23, 2014, from <http://www.consejo-discapacidades.gob.ec>.

Damico, J. S., Müller, N., & Ball, M. J. (2010). *The handbook of language and speech disorders*. Singapore: Wiley-Blackwell.

Frades, I., & Rune, M. (2010). Overview on techniques in cluster analysis. *Bioinformatics Methods in Clinical Research*, 81–107.

Hariharan, M., Chee, L. S., Chia Ai, O., & Yaacob, S. (2012). Classification of speech dysfluencies using LPC based parameterization techniques. *Journal of Medical Systems, 36*(3), 1821–1830.

Hogrefe, K., Ziegler, W., Wiesmayer, S., Weidinger, N., & Goldenberg, G. (2013). The actual and potential use of gestures for communication in aphasia. *Aphasiology, 27*(9), 1070–1089.

Hussin, M. F., Kamel, M. S., & Nagi, M. H. (2004). An efficient two-level SOMART document clustering through dimensionality reduction. *Lecture Notes in Computer Science, 3316*, 158–165.

Khalifa, A. A., Haranczyk, M., & Holliday, J. (2009). Comparison of nonbinary similarity coefficients for similarity searching, clustering and compound selection. *Journal of Chemical Information and Modeling*, 1193–1201.

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

Kostoulas, T., Mporas, I., Kocsis, O., Ganchev, T., Katsaounos, N., Santamaría, J. J., et al. (2012). Affective speech interface in serious games for supporting therapy of mental disorders. *Expert Systems with Applications, 39*(12), 11072–11079.

Lechevallier, Y., & Ciampi, A. (2007). Multilevel clustering for large databases. In J.-L. Auget, N. Balakrishnan, M. Mesbah, & G. Molenberghs (Eds.), *Advances in statistical methods for the health sciences*. Boston: Birkhäuser.

Mashima, P. A., & Doarn, C. R. (2008). Overview of telehealth activities in speech–language pathology. *Telemedicine and e-Health, 14*(10), 1101–1117.

Mustafa, M. B., Rosdi, F., Salim, S. S., & Mughal, M. U. (2015). Exploring the influence of general and specific factors on the recognition accuracy of an ASR system for dysarthric speaker. *Expert Systems with Applications, 42*(8), 3924–3932.

Nguyen, V., Phung, D., Nguyen, X., Venkatesh, S., & Hai Bui, H. (2014). Bayesian nonparametric multilevel clustering with group-level contexts. In *Proceedings of the 31st international conference on machine learning*. Beijing, China.

National Institute of Deafness and Other Communication Disorders (NIDCD) (2014). *Statistics on voice, speech and language*. Retrieved May 23, 2014, from: <https://www.nidcd.nih.gov/health/statistics/pages/vsl.aspx>.

Ooi, Y. P., Raja, M., Sung, S. C., Fung, D. S., & Koh, J. B. (2012). Application of a web-based cognitive-behavioural therapy programme for the treatment of selective mutism in Singapore: A case series study. *Singapore Medical Journal, 53*(7), 446–450.

Palmer, R., Enderby, P., Cooper, C., Latimer, N., Julious, S., & Paterson, G. (2012). Computer therapy compared with usual care for people with long-standing aphasia poststroke – A pilot randomized controlled trial. *Stroke, 43*(7), 1904–1911.

Parrilla, R. & Sierra, C. (2013). *Trastornos del lenguaje*. Boletin de la SPAO, Vol. 7.

Pennington, L., Goldbart, J., & Marshall, J. (2003). Speech and language therapy to improve the communication skills of children with cerebral palsy. *Cochrane Database of Systematic Reviews, 3*.

Petkos, G., Papadopoulos, S., & Kompatsiaris, Y. (2014). Two-level message clustering for topic detection in Twitter. In *Proceedings of the SNOW 2014 data challenge*. Seoul, Korea.

Robles-Bykbaev, V., López-Nores, M., Pazos-Arias, J. J., García-Duque, J., & Ochoa-Zambrano, J. (2014). *Modelling domain knowledge of speech and language therapy with an OWL ontology and OpenEHR archetypes*. University of Vigo Technical Report UVI-SSI-14-008.

Saz, O., Yin, S., Lleida, E., Rose, R., Vaquero, C., & Rodriguez, W. (2009). Tools and technologies for computer-aided speech and language therapy. *Speech Communication, 51*(10), 948–967.

Schipor, O. A., Pentiuc, S. G., & Schipor, M. D. (2010). Improving computer-based speech therapy using a fuzzy expert system. *Computing & Informatics, 29*(2).

Schipor, O. A., Pentiuc, S. G., & Schipor, M. D. (2012). Automatic assessment of pronunciation quality of children within assisted speech therapy. *Electronics & Electrical Engineering*, 15–18.

Sekine, K., & Rose, M. L. (2013). The relationship of aphasia type and gesture production in people with aphasia. *American Journal of Speech-Language Pathology, 22*(4), 662–672.

Teodoro, G., Martin, N., Keshner, E., Shi, J. Y., & Rudnicky, A. (2013). Virtual clinicians for the treatment of aphasia and speech disorders. In *10th International conference on virtual rehabilitation (ICVR)* (pp. 158–159).

Utianski, R. L., Sandoval, S., Lehrer, N., Berisha, V., & Liss, J. (2013). Speech assist: An augmentative tool for practice in speech–language pathology. *Journal of the Acoustical Society of America, 134*(5).

Verikas, A., Bacauskiene, M., Gelzinis, A., Vaiciukynas, E., & Uloza, V. (2012). Questionnaire-versus voice-based screening for laryngeal disorders. *Expert Systems with Applications, 39*(6), 6254–6262.

Yeh, Y. L., Hou, T. H., & Chang, W. Y. (2012). An intelligent model for the classification of children's occupational therapy problems. *Expert Systems with Applications, 39*(5), 5233–5242.

Zhang, H., & Chow, T. W. S. (2012). A multi-level matching method with hybrid similarity for document retrieval. *Expert Systems with Applications, 39*(3), 2710–2719.