Decision Support Systems and Education
J. Mantas et al. (Eds.)

doi:10.3233/978-1-61499-921-8-185

A Speech Sound Disorder Screening System Database Structure

Emilian Erman MAHMUT[a,1] and Vasile STOICU-TIVADAR[a]
[a] *Politehnica University of Timisoara, Department of Automation and Applied Informatics*

Abstract. This paper makes a brief review of several database structures of Computer-Based Speech Therapy (CBST) systems and solutions and describes the screening method, an experimental study conducted to validate the screening algorithm and a database structure for the Information Entropy-Based Sound Speech Disorder (SSD) Screening System aimed at by our research project. The final part briefly presents the essential design criteria and further development.

Keywords. Speech Sound Disorders, Dyslalia, Computer-based Speech Therapy Tools, Information Entropy, Screening System.

1. Introduction

A series of complex factors make the field of computer-aided SSD detection highly challenging: transdisciplinarity (speech pathology, neuroscience, psychology, acoustics, linguistics, statistics and probability, IT), the subsequent lack of consistent, cross-refereanceable terminology, and the prevalence of SSD therapy-oriented and language-dependent automated solutions as compared to the number of available automated SSD screening tools. Moreover, especially with young subjects, the issue of personal data confidentiality, when combined with poor awareness of the potentially devastating effects of SSDs undetected in due time on the educational process of primary school children, often frustratingly limits the access to sample materials for research purposes.

The etiology of Speech Sound Disorders identifies a variety of causes ranging from developmental disorders (language disorders, motor disorders, autism spectrum disorders and learning disorders), hearing impairment and genetic syndromes (trisomy 21 better known as the Down syndrome) all the way to brain damage (cerebral palsy or physical trauma to the brain caused by skull injuries). To give but one example of how intricate the neurobiological mechanisms supporting articulate speech can be, [1] provides a detailed description of phonological encoding in relation to the Levelt psycholinguistic model of language production, interestingly noticing that children affected by SSDs and their typically developing peers achieve similar results in a picture-naming task, which indicates that there is no causal relationship between the phonological encoding ability and flawed pronunciation.

[1] Corresponding Author, Emilian Erman MAHMUT, Automation and Applied Informatics Department, Politehnica University Timisoara, 2 Vasile Parvan Blvd., 300223 Timisoara, Romania; E-mail: emilian.mahmut@aut.upt.ro.

Given the ontogenetic development of human speech and, in particular, the ontogeny of speech sounds in children it becomes obvious why most speech pathologists agree that therapeutic intervention no later than the early school age (5-7) is of paramount importance [2][3]. In terms of frequency of occurrence, dyslalic disorders hold the central place among SSDs and are associated with the Distortion, Substitution, Omission or Inversion (DOSI Score) of phonemes in pronunciation. Furthermore, the SSDs' global trend is ascending, with 25-30% of dyslalic children being also affected by dyslexia-dysgraphia, and dyslalia along with rhythm disorders (stuttering) are associated to left-handedness and right-brain dominance [3].

This paper's objective is to propose an original database structure for the Entropy-based SSD Screening System which is the ultimate objective of our research project.

2. Review of existing CBST Database Structures

The TERAPERS Solution [4] is a complex CBST for the personalized therapy of early-school age dyslalic Romanian children. The *"Complex Examination"* module which performs an initial evaluation of the subject consists in a series of standardized tests to be administered to children by Speech Language Pathologists (SLPs) in educational institutions. The subjects' data, including the initial evaluation data, are stored in a relational database which is accessed during therapy to assess the subject's progress and suggest therapy exercises. The *Logomon* application is a fix part of the system installed on each SLP's office computer, which contains a local database and represents a node of the whole distributed database which may be reached via an internal network or online. The authors claim that this configuration with a unique database management system covering all the nodes which use the same application (*Logomon*) provides faster access to the data and reduces the operation costs. The relational database is managed by an Oracle management system, which raises the issue of cost-efficiency. Paper [5] describes ISLand, an intelligent system supporting the language development of children, organized in layers and modules. The Access Interface is connected to the Information Management layer which includes 6 modules. The Information Management layer is connected to a Knowledge Database made up of 4 distinct linguistic levels (Phonological, Semantic, Morphosyntactic and Pragmatic) also accessed by a Decision-Making block made up of 3 modules (Statistical Validation, Reports and Monitoring and Therapy Sessions Planner). The system makes extensive use of games and its Access Interface may be used by different user profiles (children, teachers, family) from desktop and mobile devices. [6] describes a Speech Therapy system made up of 3 main blocks (Server, Game and Therapist Portal) designed for English speaking children aged 5 to 12. The Central recording Server, placed between the Game and the Therapist Portal, is written in Node JS and stores data in a MongoDB instance. The server contains 5 modules: Therapist, Storage, Child, File Encoder and Recording. A Server API connects *Kimbee* (the Game) to the central server. Three data models are stored in the server: recordings (audio files processed in the Speech Processing module), children (subject profiles) and therapists (essentially operator login credentials). This solution also uses gamification.

Another interesting approach based on an Item Response Theory (IRT) model applied to a standardized children speech production test is presented in paper [7].

3. Entropy-Based SSD Screening Method and Experimental Study

In [8] we described the Information Entropy-based SSD screening method used in our experimental study for the assessment of the similarity between a standard sound pattern (SLP's pronunciation) and a sample sound pattern (subject's pronunciation). The algorithm that we used discriminates mispronunciations by segmenting the sound patterns (amplitude crests and troughs), by calculating the information content (entropy) of each sound pattern and comparing the results, which are then displayed, by association of the information value with color bands, as an isometric diagram.

As experimental study, 30 primary school subjects aged 6 affected by dyslalia (rhotacism) were given a speech therapy worksheet and were asked to pronounce several Romanian words containing the /r/ phoneme in the initial, median and final position. The audio signal was recorded with an Olympus LSP1 Linear PCM Recorder (44.1 kHz/16 bit) and converted to digital signal with the open-source software Audacity 1.3 Beta (Unicode) before being fed into the algorithm. The experimental study we performed was meant to provide an initial validation for the algorithm. The results of the analysis of the audio samples allowed us to ascertain that the algorithm identifies correctly the differences between the pronunciation of the standard pattern and the pronunciation of the sample pattern.

4. Database Structure Proposal

In pursuit of the research thread discussed in [8] our next objective is to develop an integrated system that would allow the management of SLPs, subjects (children tested for dyslalic disorders), of the audio samples and of the analysis algorithms, with an automated processing of the audio recordings and presentation of the results (isometric diagram). The general diagram of this Screening System is shown in Figure 1 below.

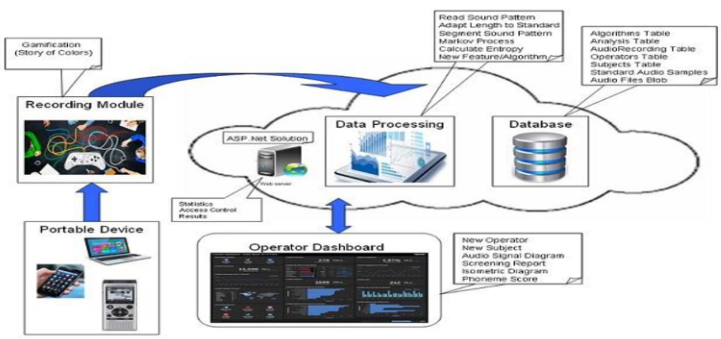

Figure 1. General Diagram of the Screening System

The screening subjects data and the screening results are stored in a relational database. The main tables in our database (Figure 2 below) are: Algorithms, Analysis, Subjects, Operators and AudioRecordings. The Analysis table makes a connection between the Operators table and the Subjects table in order to ensure the control of the database access level. The fields in this table make the unique connection between each Operator (SLP) and his/her subjects.

The name of the audio file for each subject will also be composed from this table ("Operator_ID + Subject_CNP + AudioRecording_ID.file + Algorithm_ID"). The

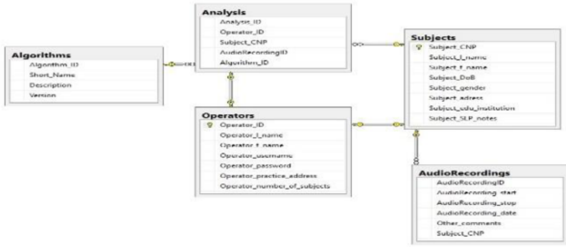

Figure 2. Main Table Relations

Operators table contains data related to each operator (SLP), while the Subjects table contains data on each subject. These two tables are connected through the Operator_ID and Subject_CNP fields so that more operators may not have the same subject.

The AudioRecordings table is related to the Subjects table which contains data on the audio recordings of each subject. The Algorithms Table fields store all the screening algorithm-related data employed in the research activity. This table is related to the Analysis Table in order to allow the identification of the selected screening algorithm/algorithm version. The main tables in the relational database are implemented in SQL Server Management Studio. We opted for this solution keeping in mind cost-efficiency. An encryption algorithm will be used to tackle the Subject personal data confidentiality issue. The data in the database is accessed as per the sequence diagram shown in Figure 3 below, which describes the operational steps of our SSD Screening System designed to also be used as a research infrastructure.

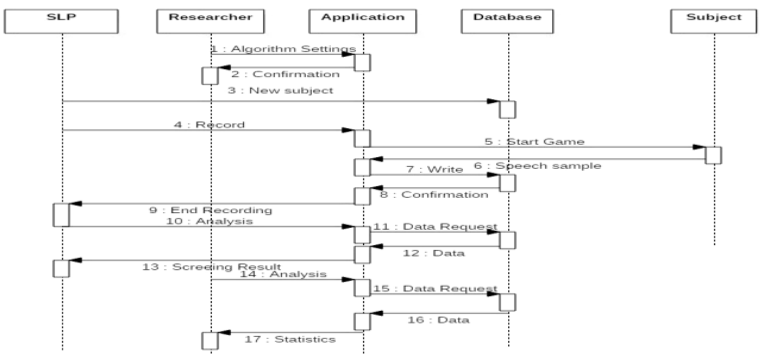

Figure 3. Sequence Diagram

The initial step allows the Researcher to configure the algorithm (new feature extraction, new algorithm, refinement of the same algorithm) in the application and to receive confirmation. The Operator (SLP) enters a New Subject in the database and then initiates the recording (Record), which triggers the start of the game (Start Game). The subject plays the game and the speech sample is acquired in the Application which sends it to the database (Write). The newly acquired speech sample storage in the database is confirmed (Confirmation) and the recording is stopped (End Recording). The following stages are dedicated to the Analysis of the speech samples performed both by the SLP (for own subjects) and by the Researcher (all existing subjects). In the final steps the results are displayed to the desktop or mobile devices used by the SLP and by the Researcher.

5. Conclusions & Discussion

The automated SSD Screening system reduces the large amount of paperwork needed to perform the screening in the traditional way, using pen and paper, and a lot of speech therapy forms and materials. It also approaches the issue of shortage of Speech Therapy specialists in rural/remote areas. By means of gamification, it also boosts the motivation of the test subjects. The system would also offer the possibility to easily access statistics by different criteria of interest for research purposes. The SSD Screening System and its database structure described above were designed keeping in mind several essential criteria such as: time-efficiency and mobility, cost-efficiency, modularity and flexibility, i.e. allowing the researcher to add new features or new algorithms. A limitation of the current algorithm is that it uses the SLP's (adult) pronunciation as the standard pattern to compare the subject's (child) pronunciation with. Future work will focus on further refining the algorithm so as to include data related to the position of the target phoneme within the word, data related to the DOSI Score and to relevant human auditory perception time thresholds, which could offer the possibility to identify the sub-types of dyslalic disorder affecting the screening subjects.

Work on the system is in progress, we intend to create the possibility to also use other algorithms for comparison and validation purposes, and then focus on the evaluation of the medical impact of the system.

References

[1] B. Munson, M.O.P. Krause, Phonological encoding in Speech Sound Disorder: Evidence from a Cross-Modal Priming Experiment, *International Journal of Language and Communication Disorders* **52** (3) (2017), 285-300 Last accessed on April 12, 2018 at: https://www.researchgate.net/publication/303288411_Phonological_Encoding_in_Speech_Sound_Disorder_Evidence_from_a_Cross-Modal_Priming_Experiment, https://doi.org/10.1111/1460-6984.12271.

[2] G. Kokkalia, A. Drigas, A. Economou, P. Roussos, Screening Tools for Kindergarten Children, *International Journal of Engineering and Science* (*iJES)* **5** (4) (2017), 76-85. Last accessed on April 12, 2018 at: http://online-journals.org/index.php/i-jes/article/view/8013

[3] E. Verza, *Tratat de logopedie*, Editura Semne, București, 2000. Capitolul VI. Tulburările de pronunție de tip dislalic; p. 124-147.

[4] M. Danubianu, S. Gh. Pentiuc, O. A. Schipor, M. Nestor, I. Ungureanu, D. M. Schipor, TERAPERS – Intelligent Solution for Personalized Therapy of Speech Disorders, *International Journal on Advances in Life Sciences* **1** (1), 2009. Last accessed on April 24, 2018 at: https://www.researchgate.net/publication/228852446

[5] E. Bernal-Merchán, P. Suquilanda-Cuesta, C. Espinoza-Fajardo, A. León-Pesántez, Y. Robles-Bykbaev, V. Robles-Bykbaev, D. Quisi-Peralta, ISLanD: An Informatics Intelligent System to Support the Language Development of Children from 4 to 5 years, *2017 IEEE International Autumn Meeting on Power, Electronics and Computing* (ROPEC 2017), Ixtapa, Mexico. Last accessed on April 24, 2018 at: https://ieeexplore.ieee.org/abstract/document/8261597/

[6] R. Drapeau, N. Huynh, A. Nech, *KIMBEE: A Speech Therapy Application for Children*. Last accessed on April 24 2018 at: https://courses.cs.washington.edu/courses/cse481i/16au/ReportArchive/Kimbee.pdf

[7] T. Brackenbury, M.J. Zickar, B. Munson, H.L. Storkel, Applying Item Response Theory to the Development of a Screening Adaptation of the Goldman-Fristoe Test of Articulation-Second Edition, *Journal of Speech, Language, and Hearing Research (ASHA)* **60** (9) (2017), 2672-2679. doi: 10.1044/2017_JSLHR-L-16-0392.

[8] E.E. Mahmut, M. Della Ventura, V. Stoicu-Tivadar, An Entropy-based Computer Model for the Measurement of Phonetic Similarity: Dyslalia Screening in Early School-Age Children, *Applied Medical Information Journal (UMF Cluj-Napoca)* **40** (1-2) (2018), 15-23. Last accessed on July 6th, 2018 at: http://ami.info.umfcluj.ro/index.php/AMI/article/view/625.